



Universidade Federal Rural de Pernambuco
Pró-Reitoria de Pós Graduação
Programa de Pós-Graduação em Informática Aplicada (PPGIA)

EDITAL DE SELEÇÃO PARA INGRESSO REGULAR NO PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA APLICADA (MESTRADO ACADÊMICO) EM 2027.1

1. Processo seletivo

1.1 As inscrições deverão ser realizadas no período de **21 de Setembro de 2026 a 19 de Outubro de 2026** conforme [calendário](#) publicado na página da Pró-Reitoria de Pós-graduação (PRPG-UFRPE), **exclusivamente online através do [SIGAA - Sistema Integrado de Gestão de Atividades Acadêmicas \(UFRPE\)](#)**.

1.2 Toda a **documentação deve ser entregue exclusivamente online através do [SIGAA](#)** e deverá seguir o disposto no item 3 deste edital, sendo indeferidas as inscrições que não seguirem estritamente o disposto no edital e no Manual do Candidato da PRPG-UFRPE.

1.3 O número de vagas **para a seleção 2027.1 do PPGIA/UFRPE é de 62 (sessenta e duas) vagas, distribuídas da seguinte forma: 01 (uma) vaga** exclusiva para servidores da UFRPE; 20% ou seja **13 (treze) vagas** destinadas ao processo de cotas para ações afirmativas; 8% ou seja **5 (cinco) vagas** destinadas a pessoas com deficiência; e **43 (quarenta e três) vagas** para ampla concorrência.

1.4 Havendo disponibilidade de vagas, poderão ser convocados(as) candidatos(as) aprovados(as), mas não classificados(as), obedecendo à ordem de classificação final.

2. Pré-requisitos para inscrição

2.1 O(A) candidato(a) deve se encaixar em uma das duas situações abaixo:

- Ter concluído **curso de graduação** em instituição reconhecida pelo Ministério da Educação, **na área de Computação e afins** (especificadas no item 4.7.1 deste edital); ou
- Ter concluído **curso de graduação** em instituição reconhecida pelo Ministério da Educação, em curso diferente dos especificados no item 4.7.1 deste edital, e ter comprovadamente cursado no mínimo **120 horas/aula** de cursos ou disciplinas que **envolvam programação** de computadores.

2.2 O(A) candidato(a) deve enviar toda a documentação exigida no formulário de inscrição.

2.3 O currículo do(a) candidato(a) deve estar **cadastrado na Plataforma Lattes** do CNPq. (<https://lattes.cnpq.br/>).

3. Documentação necessária

3.1 O(A) candidato(a) deverá preencher o **Formulário de Intenções de realização do curso de mestrado no PPGIA/UFRPE** diretamente no [SIGAA](#). O conteúdo do formulário de intenções pode ser consultado no Anexo I deste edital.

3.1.1 O(A) candidato(a) poderá indicar **até 3 (três) projetos de pesquisa** em que tenha interesse, apresentando uma lista de projetos de forma priorizada, onde o primeiro projeto da lista será considerado o projeto de maior interesse e o último, o de menor interesse do candidato (Formulário de Intenções – Anexo I).

3.1.2 Se o(a) candidato(a) apresentar, como indicação de projetos de pesquisa, uma lista com mais de **3 (três)** escolhas, todos os projetos a partir do **quarto** serão desconsiderados.

3.1.3 Os projetos de pesquisa propostos pelos docentes para a seleção de 2027 podem ser consultados no **Anexo II - Relação de Projetos de Pesquisa PPGIA/UFRPE**.

3.2 O(A) candidato(a) deve enviar a seguinte documentação através do [SIGAA](#):

- Documento de identidade e Cadastro de Pessoa Física;
- Título de eleitor;
- Comprovante de votação da última eleição ou Certidão de quitação eleitoral;
- Certificado de reservista (somente para candidatos do sexo masculino);
- Comprovante de pagamento ou isenção da taxa de inscrição (ver Manual do Candidato da PRPG);
- Diploma de graduação;
- Histórico escolar da graduação;
- Certificados ou declarações comprovando 120 horas/aula de cursos ou disciplinas que envolvam a programação de computadores (apenas para quem não tem graduação na área de Computação e afins - ver item 2.1);
- Currículo Lattes;
- Documentação comprobatória do Currículo Lattes (ver item 3.3);
- Termo de responsabilidade de candidatura exclusiva (ver modelo no Manual do Candidato da PRPG);
- Declaração de concordância da chefia (somente para servidores - ver modelo no Manual do Candidato da PRPG);
- Documentação para vagas de ações afirmativas (somente para candidatos concorrendo no sistema de cotas - ver Manual do Candidato da PRPG);
- Formulário de requerimento para atendimento especializado (somente para pessoas com deficiência - ver modelo no Manual do Candidato da PRPG).

3.3 A **documentação comprobatória do Currículo Lattes** (certificados de participação em eventos, comprovações de artigos publicados, certificados de iniciação científica e quaisquer outros comprovantes relevantes para a avaliação do(a) candidato(a) conforme descrito na seção 4.7 deste edital) deve ser **organizada na mesma ordem em que aparece no currículo e enviada em um único arquivo**.

3.4 O(a) candidato(a) assume total responsabilidade pelas informações prestadas, arcando com as consequências de eventuais erros no preenchimento, envio do formulário de inscrição e documentação.

3.5 Não é necessário autenticar a documentação. Portanto, será considerado que o(a) candidato(a) apresentará documentos que sejam a expressão da verdade.

4. Processo de seleção e classificação

4.1 O processo seletivo é conduzido por uma comissão especial de seleção, designada através de portaria emitida pela coordenação de curso.

4.2 O processo seletivo 2027 para o PPGIA/UFRPE será dado por ranqueamento das notas finais dos candidatos e candidatas.

4.2.1 O total de vagas oferecidas neste processo seletivo corresponde à soma das vagas dos projetos de pesquisa dos docentes do programa, listados no Anexo II.

4.2.2 Para cada projeto de pesquisa poderão ser selecionados candidatos de acordo com o número de vagas estabelecido para o referido projeto.

4.2.3 Uma vez que cada projeto de pesquisa apresenta uma ou mais vagas, o(s) candidato(s) mais bem classificado(s) será(ão) selecionado(s) para tais vagas. Caso o(a) candidato(a) seja selecionado(a) em mais de um projeto de pesquisa, ele(a) será classificado(a) para o projeto de pesquisa de maior prioridade segundo a sua escolha, apresentada no Formulário de Intenções.

4.2.4 Caso algum projeto de pesquisa não tenha nenhum(a) candidato(a) inscrito(a) ou aprovado(a), as vagas poderão ser ocupadas por candidatos(as) aprovados(as) mas não classificados(as) nos projetos de sua preferência.

4.3 A vaga exclusiva para servidor(a) da UFRPE será alocada pelo(a) candidato(a) que seja aprovado(a) e alcance melhor ranqueamento. Casos de empate entre os servidores da UFRPE serão resolvidos conforme o item 5.3.

4.3.1 Uma vez o(a) candidato(a) declarando-se servidor(a) da UFRPE, estará concorrendo exclusivamente à vaga de servidor(a), não concorrendo com os demais candidatos(as) não servidores(as).

4.4 As vagas exclusivas para cotas serão alocadas pelos candidatos(as) aprovados(as) que se enquadrem no processo de cotas que alcancem o melhor ranqueamento. Casos de empate entre candidatos(as) cotistas serão resolvidos conforme o item 5.3.

4.5 Após conferência da documentação, as inscrições homologadas serão encaminhadas para avaliação e pontuação, em duas etapas: avaliação do *curriculum vitae* (realizada pela comissão de seleção); e avaliação do formulário de intenções no contexto de cada projeto (realizada pelo(a) docente responsável por cada projeto).

4.6 A nota final do(a) candidato(a) é a maior nota obtida dentre as notas associadas aos projetos de pesquisa selecionados pelo(a) candidato(a) (N_{FP}). A nota associada a cada projeto de pesquisa (P) de interesse do(a) candidato(a) é calculada conforme a fórmula abaixo (onde $0,0 \leq N_{FP} \leq 10,0$):

$$N_{FP} = \text{MIN}(10,0, N_{CV} \times F_P)$$

onde:

- **MIN** é uma função definida como $\text{MIN}:R \times R \rightarrow R$, que recebe dois valores reais e retorna o menor valor entre eles;
- N_{CV} é a nota da avaliação do *curriculum vitae*;
- F_P é o fator atribuído para o projeto na avaliação do formulário de intenções.

4.7 A nota do *curriculum vitae* (N_{CV}) é atribuída para cada candidato(a), com $0,0 \leq N_{CV} \leq 10,0$, independentemente de seus projetos de interesse, conforme as fórmulas abaixo:

$$N_{CV} = 10,0 \times (P_{CV}/20,0)^{1/3}, \text{ com}$$

$$P_{CV} = N_H \times (F_{CE} + F_{PC} + F_{PD} + F_D)$$

onde:

- P_{CV} é a **pontuação do *curriculum vitae***, onde $0,0 \leq P_{CV} \leq 20,0$;
- N_H é a **nota do histórico**, referente à média geral da graduação do candidato;
- F_{CE} , F_{PC} , F_{PD} , F_D são fatores relativos ao **curso de egresso**, **produção científica**, **experiência em pesquisa e desenvolvimento (P&D)**; e **experiência em docência** do(a) candidato(a), respectivamente;
- A apresentação da documentação comprobatória válida é **condição necessária para a pontuação** dos itens do *curriculum vitae*.

4.7.1 O **Fator de Curso de Egresso (F_{CE})** assume o valor **0,1** caso o candidato seja egresso de um dos seguintes cursos: *Bacharelado em Ciência da Computação*, *Bacharelado em Sistemas de Informação*, *Bacharelado em Engenharia da Computação*, *Bacharelado em Tecnologia da Informação* e *Licenciatura em Computação*. Caso contrário, este fator assume o valor **0,0**.

4.7.2 O **Fator de Produção Científica (F_{PC})**, com $0,0 \leq F_{PC} \leq 0,4$, é calculado conforme a fórmula:

$$F_{PC} = \text{MIN}(0,4, 0,15 \times Q_{CL} + 0,15 \times Q_{AP} + 0,1 \times Q_{AC} + 0,05 \times Q_{OP})$$

onde,

- **MIN** é uma função definida como $\text{MIN}:R \times R \rightarrow R$, que recebe dois valores reais e retorna o menor valor entre eles;
- Q_{CL} é a quantidade de capítulos de livros publicados;
- Q_{AP} é a quantidade de artigos publicados em periódicos;
- Q_{AC} é a quantidade de artigos publicados em anais de conferências;
- Q_{OP} é a quantidade de outras publicações;
- F_{PC} é o Fator de Produção Científica;

- A comissão de seleção pode desconsiderar itens que não se enquadrem na área do programa, ou não atinjam critérios de qualidade mínimos (por exemplo, fator de impacto do veículo de publicação, ou classificação segundo o Qualis CAPES).

4.7.3 O **Fator de Experiência em P&D (F_{PD})**, com $0,0 \leq F_{PD} \leq 0,3$, é calculado conforme a fórmula:

$$F_{PD} = \text{MIN}(0,3, 0,15 \times Q_p + 0,02 \times Q_e)$$

onde,

- **MIN** é uma função definida como $\text{MIN}:R \times R \rightarrow R$, que recebe dois valores reais e retorna o menor valor entre eles;
- **Q_p** é a quantidade total de anos, arredondando para baixo, de que o(a) candidato(a) participou como bolsista ou voluntário em programas voltados para pesquisa científica e tecnológica, docência ou extensão (PIBIC, PIC, PIBITI, PIBID e outros programas institucionais);
- **Q_e** é a quantidade de eventos científicos de que o(a) candidato(a) participou, organizou ou onde apresentou trabalhos;
- A comissão de seleção pode desconsiderar itens que não se enquadrem na área do programa, ou não atinjam critérios de qualidade mínimos (por exemplo, classificação Qualis CAPES dos eventos, ou reconhecimento da instituição organizadora).

4.7.4 O **Fator de Experiência em Docência (F_D)**, com $0,0 \leq F_D \leq 0,2$, é calculado conforme a fórmula:

$$F_D = \text{MIN}(0,2, 0,1 \times Q_{ES} + 0,05 \times Q_{EO} + 0,02 \times Q_M)$$

onde,

- **MIN** é uma função definida como $\text{MIN}:R \times R \rightarrow R$, que recebe dois valores reais e retorna o menor valor entre eles;
- **Q_{ES}** é a quantidade total de anos, arredondando para baixo, em que o(a) candidato(a) lecionou no ensino superior;
- **Q_{EO}** é a quantidade total de anos, arredondando para baixo, em que o(a) candidato(a) lecionou em outros níveis de ensino, como fundamental, técnico ou médio;
- **Q_M** é a quantidade total de anos, arredondando para baixo, em que o(a) candidato(a) atuou como monitor(a) em instituições de ensino;
- A comissão de seleção pode desconsiderar itens que não se enquadrem na área do programa, ou não atinjam critérios de qualidade mínimos (por exemplo, atuação como professor particular ou em instituição não reconhecida pelo MEC).

4.8 O Fator do Projeto (F_p), com $0,5 \leq F_p \leq 1,5$, é um fator atribuído à adequação do(a) candidato(a) ao projeto específico, com base nas informações prestadas no Formulário de Intenções de Realização do Curso de Mestrado no PPGIA/UFRPE.

4.8.1 A análise será baseada na relevância das habilidades técnicas e pessoais relatadas

pelo(a) candidato(a) para o projeto de pesquisa;

4.8.2 O valor de pontuação para o formulário de intenções será de 0,5 a 1,5.

- 4.9** A nota final do(a) candidato(a) será utilizada para a geração do ranqueamento. Será considerado(a) **APROVADO(A)** o(a) candidato(a) cuja nota final seja maior ou igual a 6,0. Caso contrário, será considerado(a) **REPROVADO(A)**.

5. Resultado Final

- 5.1** O resultado do Processo Seletivo para o PPGIA/UFRPE será divulgado considerando cada projeto de pesquisa.

- 5.2** A distribuição de vagas se dará da seguinte forma:

5.2.1 As vagas disponíveis no projeto serão preenchidas pelos(as) candidatos(as) aprovados(as), de acordo com o ranqueamento geral, sendo o(a) candidato(a) considerado(a) CLASSIFICADO(A);

5.2.2 Caso o(a) candidato(a) seja classificado(a) em mais de um projeto, ele(a) será selecionado(a) para o projeto de maior prioridade segundo informações dadas no Formulário de Intenções de Realização do Curso de Mestrado no PPGIA/UFRPE (Anexo I).

5.2.3 Os(As) candidatos(as) aprovados(as), porém não classificados(as) nos projetos de sua escolha, serão convocados(as) para vagas remanescentes de outros projetos.

5.2.4 Os(As) candidatos(as) aprovados(as), porém não classificados(as) de acordo com o ranqueamento geral poderão ser convocados(as) caso haja desistência dos(as) candidatos(as) classificados(as), seguindo a ordem decrescente das notas finais.

- 5.3** Eventuais empates serão resolvidos de acordo com os valores de (nesta ordem): Fator do Projeto (F_p), Nota do *Curriculum Vitae* (N_{cv}), Fator de Produção Científica (F_{pc}), Fator de Experiência em P&D (F_{pd}), Fator de Experiência em Docência (F_d).

- 5.4** Os resultados serão publicados no site <https://www.ppgia.ufrpe.br/>.

- 5.5** Sobre eventuais concessões de bolsas de estudo:

5.5.1 O(A) candidato(a) aprovado(a) e classificado(a) poderá receber bolsa de estudo caso esteja habilitado(a) para este recebimento, segundo as Normas Gerais dos Programas de Pós-Graduação Stricto Sensu da UFRPE, Normas Específicas do PPGIA/UFRPE e normas gerais das Agências de Fomento à Pesquisa.

5.5.2 No PPGIA/UFRPE existem dois tipos de bolsas de estudo - nível mestrado: (a) bolsa de estudo vinculada a um projeto de pesquisa; (b) bolsa de estudo da cota do PPGIA/UFRPE.

5.5.2.1 As bolsas de estudo vinculadas a um projeto de pesquisa são de responsabilidade do(a) docente coordenador(a) do projeto e serão concedidas aos(às)

candidatos(as) aprovados(as) e classificados(as) no respectivo projeto de pesquisa, de acordo com a disponibilidade de bolsas e elegibilidade dos candidatos.

5.5.2.2 As bolsas de estudo da cota PPGIA/UFRPE, caso haja disponibilidade, são oferecidas aos(às) candidatos(as) classificados(as) e aptos(as) a recebê-las seguindo as regras estabelecidas na Minuta de Gestão de Bolsas do PPGIA (disponível no site do programa) e a resolução da UFRPE sobre concessão de bolsas.

5.6 Fica assegurado ao(à) candidato(a) o direito de recorrer do resultado final da seleção, no prazo de até 03 (três) dias de sua divulgação, conforme definido no cronograma de seleção. O recurso deve ser obrigatoriamente aberto por meio do [SIGAA](#).

6. Cronograma

6.1 A seleção para o PPGIA/UFRPE consistirá das seguintes etapas:

Etapas do processo seletivo do mestrado	Datas
Período de Inscrições (via SIGAA)	21/Set/2026 a 19/Out/2026
Período de análise das inscrições para fins de homologação	20/Out/2026 a 27/Out/2026
Homologação das inscrições	28/Out/2026
Período recursal referente à homologação das inscrições	29/Out/2026 a 03/Nov/2026
Publicação das respostas aos recursos referentes à homologação das inscrições	05/Nov/2026
Período de avaliação da documentação das inscrições homologadas	06/Nov/2026 a 19/Nov/2026
Publicação da consolidação parcial do processo seletivo	23/Nov/2026
Período recursal referente à consolidação parcial	24/Nov/2026 a 26/Nov/2026
Publicação das respostas aos recursos referentes à consolidação parcial	27/Nov/2026
Publicação da consolidação final do processo seletivo	01/Dez/2026
Período recursal referente à consolidação final	02/Dez/2026 a 04/Dez/2026

Publicação das respostas aos recursos referentes à consolidação final	07/Dez/2026
Publicação dos resultados finais	09/Dez/2026
Período de matrículas	01/Mar/2027 a 03/Mar/2027
Início das aulas	08/Mar/2027

6.2 Os resultados das etapas eliminatórias serão publicados no SIGAA e no endereço <https://www.ppgia.ufrpe.br/>.

7. Disposições Gerais

7.1 A realização da inscrição implica a plena concordância do candidato ao edital e às normas complementares.

7.2 A Comissão Especial de Seleção decidirá os casos omissos.

ANEXO I - Formulário de Intenções de Realização do Curso de Mestrado no PPGIA/UFRPE
(Este formulário está disponível no SIGAA e deve ser preenchido diretamente no sistema.)

1) Você está concorrendo às vagas destinadas a cotas?

☐ Sim

☐ Não

2) Você é servidora(a) da UFRPE?

☐ Sim

☐ Não

3) Qual(is) curso(s) de graduação você fez?

4) Em que ano/semestre você concluiu sua graduação?

5) Em qual instituição você cursou a sua graduação?

6) No seu curso de mestrado, você pretende ter dedicação:

☐ Exclusiva

☐ Parcial

7) Escreva em ordem de prioridade (da primeira para a terceira) os números e títulos de três projetos (dentre os listados no edital) dos quais você gostaria de participar, desenvolvendo sua pesquisa de mestrado:

8) Redija um texto abordando as questões que você considera mais relevantes nos projetos priorizados e explique como suas habilidades técnicas e pessoais podem contribuir para o sucesso da pesquisa (sugerimos redigir o texto em um editor, e colar no campo abaixo, para evitar perda dos dados em razão de eventuais erros do sistema). (Número máximo de Caracteres: 9999)

9) Declaração de veracidade

☐ Declaro que as informações preenchidas neste formulário refletem a expressão da verdade.

☐ Não quero fazer esta declaração (essa alternativa invalida a sua participação no processo seletivo)

ANEXO II - RELAÇÃO DE PROJETOS DE PESQUISA PPGIA/UFRPE 2027.1

PROJETO 1

Orientador: André Câmara Alves do Nascimento

Título: Modelos de Fundação para Adaptação Automática de Conteúdos e Aplicações Educacionais a Sistemas de Comunicação Aumentativa e Alternativa

Vagas: 3

Resumo:

Estudantes com Necessidades Complexas de Comunicação dependem de Sistemas de Comunicação Aumentativa e Alternativa (CAA), tipicamente organizados em pranchas de pictogramas apoiadas em vocabulários controlados, como ARASAAC, Widgit ou PCS. O material didático da escola regular, entretanto, é produzido em texto corrido e imagens não adaptadas, e sua conversão para representação pictográfica permanece um processo manual, caro e dependente de profissionais especializados escassos, o que restringe severamente o acesso desses estudantes ao currículo. Large Language Models (LLMs) vêm sendo aplicados intensamente nas áreas de educação, mais especificamente como tutores, no apoio da produção de material didático e feedback automático, bem como na comunicação assistiva, na forma de produção de conteúdos focados em comunicação. Modelos fundacionais multimodais abrem uma etapa seguinte: converter automaticamente objetos de aprendizagem completos em representações pictográficas coerentes, linguisticamente simplificadas e ancoradas em vocabulário controlado. Este projeto investiga esse pipeline em quatro frentes: (i) simplificação e telegrafização de textos educacionais guiada por modelos de linguagem, com controle explícito de nível de leitura e preservação do conteúdo curricular; (ii) ancoragem semântica de texto para pictograma sobre corpora como o ARASAAC, tratando polissemia, ambiguidade e lacunas de vocabulário como um problema de recuperação e reordenação com representações multimodais; (iii) geração condicionada de pictogramas para conceitos ausentes, com consistência de estilo visual e adequação ao contexto cultural brasileiro; e (iv) personalização ao perfil linguístico e cognitivo do usuário e ao contexto de sala de aula. A avaliação combinará métricas automáticas com estudos junto a fonoaudiólogos e professores de Atendimento Educacional Especializado. Riscos de alucinação, viés cultural, opacidade e perda de agência comunicativa do usuário serão tratados como objeto explícito de investigação, e não como externalidade, alinhando o projeto às diretrizes de educação inclusiva e às boas práticas de IA responsável.

Referências:

Roberto Pereira, Ticianne Darin, and Milene Selbach Silveira. 2024. GrandIHC-BR: Grand Research Challenges in Human-Computer Interaction in Brazil for 2025-2035 * In Proceedings of the XXIII Brazilian Symposium on Human Factors in Computing Systems (IHC '24). Association for Computing Machinery, New York, NY, USA, Article 16, 1–24. <https://doi.org/10.1145/3702038.3702061>

Pereira, J. A.; Macêdo, D.; Zanchettin, C.; Oliveira, A. L. I.; Fidalgo, R. N. "PictoBERT: Transformers for next pictogram prediction." Expert Systems with Applications, v. 202, 117231, 2022. <https://doi.org/10.1016/j.eswa.2022.117231>

Pereira, J. A. et al. "PrAACT: Predictive Augmentative and Alternative Communication with Transformers." Expert Systems with Applications, v. 240, 122417, 2024. <https://doi.org/10.1016/j.eswa.2023.122417>

Di Paola, A.; Muraro, S.; Marinelli, R.; Pilato, C. "Foundation Models in Augmentative and Alternative Communication: Opportunities and Challenges." arXiv preprint arXiv:2401.08866, 2024.

Valencia, S.; Cave, R.; Kallarackal, K.; Seaver, K.; Terry, M.; Kane, S. K. "'The less I type, the better': How AI Language Models Can Enhance or Impede Communication for AAC Users." In: Proceedings of the 2023 CHI

Conference on Human Factors in Computing Systems (CHI '23), ACM, 2023. <https://doi.org/10.1145/3544548.3581560>

"Future technologies in alternative and augmented communication: a scoping review of innovations." *Frontiers in Communication*, 2025. <https://doi.org/10.3389/fcomm.2025.1607531>

PROJETO 2

Orientador: André Câmara Alves do Nascimento

Título: Aprendizagem de Representação e Graph Neural Networks para a Descoberta Acelerada de Materiais Eletrocatalíticos

Vagas: 1

Resumo:

A produção de hidrogênio verde por eletrólise da água depende criticamente de eletrocatalisadores eficientes e de baixo custo para a reação de evolução de oxigênio (OER), a etapa cineticamente limitante do processo. A busca experimental por materiais de alto desempenho (e.g., estruturas metal-orgânicas, materiais de alta entropia, espinélio e óxidos de metais de transição) é lenta e cara, enquanto a simulação por métodos mais precisos é muitas vezes computacionalmente proibitiva em espaços composicionais combinatórios. Revisões recentes da área (Jayabharathi et al., 2025) documentam o uso crescente de aprendizado de máquina para revelar descritores que governam o sobrepotencial e orientar o design de catalisadores, mas a maior parte desses trabalhos ainda se apoia em descritores composicionais tabulares, que descartam a estrutura tridimensional do material. Este projeto propõe evoluir essa linha para a aprendizagem de representação estrutural com Graph Neural Networks, explorando o fato de que cristais e superfícies catalíticas são grafos naturais (átomos como nós, vizinhança e ligações como arestas). As representações obtidas serão avaliadas sobre benchmarks públicos como Open Catalyst e Matbench Discovery, uma limitação reconhecida dos modelos universais atuais.

Referências:

Jayabharathi, J.; Thanikachalam, V.; Karthikeyan, B.; Sangamithirai, M.; Vijayarangan, M. "Recent progress of top-performing electrocatalytic materials for water oxidation and recent machine learning edge: An overview up to 2024." *Journal of Industrial and Engineering Chemistry*, v. 146, p. 136-175, 2025. <https://doi.org/10.1016/j.jiec.2024.11.050>

Xie, T.; Grossman, J. C. "Crystal Graph Convolutional Neural Networks for an Accurate and Interpretable Prediction of Material Properties." *Physical Review Letters*, v. 120, 145301, 2018. <https://doi.org/10.1103/PhysRevLett.120.145301>

Chen, C.; Ong, S. P. "A universal graph deep learning interatomic potential for the periodic table." *Nature Computational Science*, v. 2, p. 718-728, 2022. <https://doi.org/10.1038/s43588-022-00349-3>

Batatia, I.; Kovács, D. P.; Simm, G. N. C.; Ortner, C.; Csányi, G. "MACE: Higher Order Equivariant Message Passing Neural Networks for Fast and Accurate Force Fields." In: *Advances in Neural Information Processing Systems (NeurIPS)*, 2022.

Deng, B. et al. "CHGNet as a pretrained universal neural network potential for charge-informed atomistic modelling." *Nature Machine Intelligence*, v. 5, p. 1031-1041, 2023. <https://doi.org/10.1038/s42256-023-00716-3>

Chanussot, L. et al. "Open Catalyst 2020 (OC20) Dataset and Community Challenges." *ACS Catalysis*, v. 11, n. 10, p. 6059-6072, 2021. <https://doi.org/10.1021/acscatal.0c04525>

Loew, A. et al. "Universal machine learning interatomic potentials are ready for phonons." npj Computational Materials, 2025. <https://doi.org/10.1038/s41524-025-01650-1>

PROJETO 3

Orientador: André Câmara Alves do Nascimento

Título: Modelos Agrícolas de Grande Escala para Recomendação de Enxertia Adaptada a Microclimas: Integração de Interações, Texto e Embeddings Geoespaciais

Vagas: 1

Resumo:

A enxertia é uma tecnologia central da fruticultura perene: ela acopla dois genótipos distintos, i.e., copa e porta-enxerto, em um único organismo funcional cujas propriedades emergentes (vigor, produtividade, tolerância a seca e salinidade, qualidade de fruto) não são atribuíveis a nenhum dos parceiros isoladamente. O desempenho de uma combinação, porém, não é uma constante: a mesma dupla copa/porta-enxerto responde de forma distinta conforme o solo, o regime térmico e hídrico e o histórico de uso da área, de modo que a recomendação agrônômica é, na prática, uma função de três argumentos: copa, porta-enxerto e lugar. Os trabalhos computacionais anteriores do grupo cobriram apenas os dois primeiros: Silva et al. (2022) trataram a compatibilidade entre cultivares como predição de arestas em um grafo bipartido de interações, e Verslype et al. (2023) classificaram níveis de tolerância à seca em porta-enxertos de videira a partir de descritores fisiológicos, ambos sem representação explícita do local de cultivo. Este projeto propõe fechar essa lacuna por meio do que vem sendo chamado de modelos agrícolas de grande escala (Guo et al., 2026): arquiteturas que integram modalidades heterogêneas em uma representação única de apoio à decisão, análogas aos grandes modelos computacionais já consolidados em engenharia. A proposta articula três fontes: (i) dados de interação, na forma do grafo copa × porta-enxerto com desfechos observados em campo; (ii) conhecimento textual, e.g., literatura agrônômica, descritores de cultivares, boletins técnicos e bases de dados controladas, de modo a fornecer representações iniciais para cultivares com poucos ensaios registrados; e (iii) embeddings geoespaciais aprendidos, ex., Satellite Embedding (Brown, 2025), que resume em 64 dimensões por pixel de 10 metros a trajetória anual de condições de superfície, complementado por séries climáticas, atributos de solo e relevo. A fusão dessas modalidades permite prever o desempenho de triplas (copa, porta-enxerto, local) e enfrentar diretamente os cenários de partida a frio — cultivar nova, área nova — que limitam os modelos atuais. Serão investigadas estratégias de alinhamento entre modalidades, calibração de incerteza, transferência de áreas bem monitoradas para áreas sem histórico e protocolos de avaliação com partições espaciais e temporais, indispensáveis para evitar que a autocorrelação espacial produza estimativas otimistas. O estudo de caso é a viticultura tropical do Submédio São Francisco, e o projeto se apoia em trabalho já em andamento no grupo de consolidação de bases a partir de fontes estruturadas e não estruturadas. Espera-se produzir modelos simultaneamente preditivos e interpretáveis, capazes de sustentar recomendação de combinações sob microclimas específicos e sob cenários de mudança climática, com aplicação direta à viticultura do Vale do São Francisco e generalização para outras fruteiras enxertadas de importância econômica no Nordeste.

Referências:

Silva, T. B. R.; Verslype, N. I.; Nascimento, A. C. A.; Prudêncio, R. B. C. "Predicting Compatibility of Cultivars in Grafting Processes Using Kernel Methods and Collaborative Filtering." In: Intelligent Systems (BRACIS 2022), Lecture Notes in Computer Science, Springer, 2022. https://doi.org/10.1007/978-3-031-21686-2_42

Verslype, N. I.; Nascimento, A. C. A.; Musser, R. S.; Caldas, R. M. S.; Martins, L. S. S.; Leão, P. C. S. "Drought tolerance classification of grapevine rootstock by machine learning for the São Francisco Valley." *Smart Agricultural Technology*, v. 4, 100192, 2023. <https://doi.org/10.1016/j.atech.2023.100192>

Guo, R.; Wang, D.; Zhao, X.; Hu, H. An Overview of Large Agricultural Models: Current Status, Applications, and Future Perspectives. *Agriculture* 2026, 16, 1419. <https://doi.org/10.3390/agriculture16131419>

Brown, C. F.; Kazmierski, M. R.; Pasquarella, V. J.; Rucklidge, W. J.; Samsikova, M.; Zhang, C.; Shelhamer, E. et al. "AlphaEarth Foundations: An embedding field model for accurate and efficient global mapping from sparse label data." *arXiv preprint arXiv:2507.22291*, 2025.

Google Earth Engine; Google DeepMind. "Google Satellite Embedding V1 Annual." *Earth Engine Data Catalog*, 2025.

Ma, Y.; Shen, Y.; Swatantran, A.; Lobell, D. B. "Harvesting AlphaEarth: Benchmarking the geospatial foundation model for agricultural downstream tasks." *arXiv preprint arXiv:2601.00857*, 2026.

PROJETO 4

Orientadora: Andrêza Leite de Alencar

Título: Inteligência preditiva em saúde para a construção e fortalecimento de redes de vigilância e detecção precoce de patógenos

Vagas: 2

Resumo:

A rápida propagação de doenças infecciosas, intensificada pela mobilidade humana, pelas mudanças climáticas e pelas desigualdades socioeconômicas, representa um dos principais desafios para os sistemas contemporâneos de vigilância em saúde. Embora os sistemas tradicionais sejam fundamentais para o monitoramento epidemiológico, muitos ainda possuem caráter predominantemente reativo, baseado na identificação de casos já ocorridos. Nesse contexto, o projeto propõe o desenvolvimento e a evolução de uma plataforma de **inteligência preditiva em saúde**, capaz de integrar dados heterogêneos e identificar, de forma antecipada e geograficamente localizada, áreas de risco e possíveis rotas de propagação de patógenos. O projeto tem como objetivo transformar metodologias científicas desenvolvidas no contexto da epidemiologia, modelagem matemática e ciência de dados em uma plataforma digital de **Epidemic Intelligence**. A solução integrará dados de mobilidade, indicadores climáticos e ambientais, informações socioeconômicas e dados genômicos de patógenos. A partir dessa integração, serão desenvolvidos modelos estatísticos, matemáticos e de **Machine Learning** para identificar sinais de anomalias, estimar riscos e mapear rotas de propagação de doenças, incluindo arboviroses e vírus respiratórios. A pesquisa será desenvolvida inicialmente no território nacional, considerando suas particularidades geográficas, epidemiológicas e sociais, incluindo a necessidade de incorporar variáveis relacionadas a populações e territórios vulnerabilizados. Entre os principais resultados esperados estão o aprimoramento de modelos preditivos, a construção de um **Data Lake multissetorial**, o desenvolvimento de dashboards para apoio à tomada de decisão e a geração de mapas de risco, alertas precoces e relatórios automatizados para gestores e profissionais da saúde. A vaga de mestrado está direcionada a candidatos interessados em desenvolver pesquisas na interface entre **Ciência de Dados, Inteligência Artificial, Saúde Digital e Vigilância Epidemiológica**. Dependendo do perfil e dos interesses do estudante selecionado, poderão ser exploradas temáticas como integração e engenharia de dados, análise de mobilidade humana, modelagem matemática, algoritmos de redes, Machine Learning, análise espacial, visualização de dados e desenvolvimento de sistemas inteligentes para apoio à decisão.

Referências:

Oliveira, J. F.; Alencar, A. L.; Cunha, M. C. L. S.; et al. Human mobility patterns to inform sampling sites for early pathogen detection and routes of spread: a network modelling and validation study. *The Lancet Digital Health*, 2024.

Oliveira, J. F.; Vasconcelos, A. O.; Alencar, A. L.; et al. Balancing Human Mobility and Health Care Coverage in Sentinel Surveillance of Brazilian Indigenous Areas: Mathematical Optimization Approach. *JMIR Public Health and Surveillance*, 2025.

Edwards, A. M.; Baric, R. S.; Saphire, E. O.; et al. Stopping pandemics before they start: lessons learned from SARS-CoV-2. *Science*, 2022.

PROJETO 5

Orientador: Danilo Ricardo Barbosa de Araújo

Título: Uma Abordagem Multiobjetiva para o Projeto de Redes Subfluviais de Alta

Capacidade

Vagas: 1

Resumo:

Regiões rurais de difícil acesso podem ser beneficiadas pela adoção de tecnologias de redes subfluviais de alta capacidade. A região amazônica carece de infraestrutura de rede adequada, sobretudo por suas características geográficas, e é uma candidata à adoção deste tipo de tecnologia [1]. Programas como o Programa Amazônia Integrada e Sustentável (PAIS) [2] buscam construir uma infraestrutura de backhaul subfluvial pelos rios da região, mas ainda não há métodos sistemáticos para avaliar como esse planejamento pode ser expandido. Há iniciativas iniciais de apoio a essa expansão [1], mas uma breve revisão da literatura aponta que não há propostas que forneça um processo de otimização realista que considere diversos objetivos de projeto conflitantes, como custo, desempenho e impacto ambiental. Este projeto propõe uma abordagem multiobjetiva [3] para projetos de redes subfluviais de alta capacidade, que considere diversos objetivos conflitantes simultaneamente.

Referências:

[1] GOMES, L. F. A. ; ARAÚJO, D. R. B. . Um Framework de Apoio à Expansão da Rede de Backhaul Subfluvial na Amazônia Baseado em Dados Geoespaciais e Modelos Gravitacionais. In: XVI Conferência Nacional em Comunicações, Redes e Segurança da Informação (Encom 2026), 2026, João Pessoa.

[2] Ministério das Comunicações, "Plano Estratégico de Redes de Transporte de Telecomunicações (PERT), Brasília, 2024.

[3] SILVA, S. A. ; BASTOS-FILHO, C. J. A. ; ARAÚJO, D. R. B. ; ALMEIDA JUNIOR, R. C. ; MARTINS-FILHO, J. F. . Multi-objective Optimization for Resource Allocation in Elastic Optical Networks: Integrating Spectrum and Power Assignment. JOURNAL OF MICROWAVES, OPTOELECTRONICS AND ELECTROMAGNETIC APPLICATIONS, v. 25, p. 1-11, 2026.

PROJETO 6

Orientador: Danilo Ricardo Barbosa de Araújo

Título: Arquiteturas de Gêmeos Digitais para Monitoramento e Detecção Antecipada de Estados Críticos em Sistemas Complexos Baseados em Sensoriamento Distribuído

Vagas: 2

Resumo:

Nos últimos anos, o conceito de Digital Twin (Gêmeo Digital) emergiu como uma abordagem promissora para a representação digital contínua e dinâmica de sistemas físicos complexos, permitindo acompanhar estados operacionais, simular cenários futuros e antecipar comportamentos críticos (NELE et al., 2024). Entretanto, a literatura recente aponta que a maioria das arquiteturas atualmente disponíveis ainda apresenta limitações estruturais importantes (ABD WAHAB et al., 2024). Destacam-se a baixa autonomia na interpretação de eventos complexos, a forte dependência de especialistas humanos para análises operacionais, a limitada capacidade de detecção antecipada de anomalias em séries temporais e a baixa integração entre mecanismos preditivos e suporte automatizado à decisão (INAMDAR; VAN DRIEL; ZHANG, 2024). Este projeto propõe o desenvolvimento de uma arquitetura computacional inovadora baseada em Cognitive Digital Twins (Gêmeos Digitais Cognitivos), integrando representação digital dinâmica de sistemas físicos, aprendizado de máquina aplicado à modelagem de comportamento temporal, mecanismos inteligentes de detecção de anomalias e módulos autônomos de apoio à decisão capazes de interpretar estados operacionais complexos e produzir recomendações adaptativas em tempo real.

Referências:

ABD WAHAB, Nur Haninie et al. Systematic review of predictive maintenance and digital twin technologies challenges, opportunities, and best practices. *PeerJ Computer Science*, v.10, p. e1943, 2024.

NELE, Luigi et al. Towards the application of machine learning in digital twin technology: a multi-scale review. *Discover Applied Sciences*, v. 6, n. 10, p. 502, 2024.

INAMDAR, Adwait; VAN DRIEL, Willem Dirk; ZHANG, Guoqi. Digital twin technology—a review and its application model for prognostics and health management of microelectronics. *Electronics*, v. 13, n. 16, p. 3255, 2024.

PROJETO 7

Orientadora: Erica Teixeira Gomes de Sousa

Título: Detecção e Mitigação Proativa de Falhas em Ambientes de Computação em Nuvem

Vagas: 2

Resumo:

A crescente utilização de infraestruturas de nuvem para hospedar aplicações críticas torna a confiabilidade um requisito importante, diretamente associado ao cumprimento de acordos de nível de serviço (SLAs) e à continuidade operacional. Em pesquisas anteriores, os mecanismos de tolerância a falhas em ambientes distribuídos concentraram-se em estratégias reativas — como checkpointing, replicação e migração de tarefas após a detecção de uma falha já manifestada. Embora eficazes, tais abordagens incorrem em custos de recuperação, incluindo indisponibilidade temporária, perda de estado computacional e degradação de desempenho percebida pelo usuário final. Diante desse cenário, este trabalho situa-se na linha de pesquisa de detecção e mitigação proativa de falhas, cujo princípio fundamental é antecipar a ocorrência de eventos de falha por meio da análise contínua de indicadores de desempenho e de dependabilidade do sistema, como métricas de utilização de recursos e mensagens de log trocadas entre componentes, permitindo o acionamento de ações corretivas antes que a falha se manifeste de forma completa ou se propague para outros componentes da infraestrutura. Os trabalhos recentes apontam uma tendência de incorporação de técnicas de aprendizado de máquina e de aprendizado profundo nesse processo, tanto na etapa de detecção de anomalias quanto na de predição de falhas, com ganhos relatados em métricas como a taxa de acerto na identificação de falhas e a redução do tempo médio de detecção. Contudo, permanecem como desafios em aberto a validação desses modelos preditivos em ambientes de produção com carga de

trabalho real, a coordenação entre múltiplas instâncias virtuais durante ações de mitigação, e a análise dos custos computacionais adicionais introduzidos pelos próprios mecanismos de monitoramento e predição. Este trabalho tem como objetivo geral investigar técnicas de detecção precoce de anomalias baseadas em aprendizado profundo, com ênfase em redes LSTM e autoencoders, voltadas à identificação de falhas por exaustão de recursos em ambientes de nuvem, integrando essa detecção a mecanismos de mitigação proativa. O projeto enfatiza especificamente o desafio da coordenação entre múltiplas instâncias virtuais durante ações de mitigação, avaliando o impacto da abordagem proposta sobre a confiabilidade, disponibilidade e desempenho do sistema.

Referências

Gupta, S., Gard, G. K., Varshney, N., Rao, A. L. N., Srivastava, A. and Sharma, D., Real-Time Fault Detection and Classification in Smart Grids Using Support Vector Machines, 2024 4th International Conference on Innovative Practices in Technology and Management (ICIPTM), Noida, India, 2024, pp. 1-6.

Lokhande, S. D., Singh, H. and Manjre, B. M., "Design of an Improved Model for Real-Time Anomaly Detection in Cloud Environment Using HMS-LSTM and Attention-Augmented GNNs," 2024 International Conference on Artificial Intelligence and Quantum Computation-Based Sensor Application (ICAIQSA), Nagpur, India, 2024, pp. 1-6.

Zhao, X. and Huang, C., "Efficient Anomaly Detection Algorithm for Operational Data Based on Fuzzy Cognitive Map," 2024 3rd International Conference on Artificial Intelligence, Internet of Things and Cloud Computing Technology (AloTC), Wuhan, China, 2024, pp. 201-204

PROJETO 8

Orientadora: Erica Teixeira Gomes de Sousa

Título: Avaliação da Resiliência de Plataformas de Cidades Inteligentes em Nuvem por meio de Injeção de Falhas

Vagas: 2

Resumo:

Cidades inteligentes dependem cada vez mais de infraestruturas em nuvem para sustentar serviços essenciais à população, como semaforização, monitoramento ambiental, segurança pública, gestão energética e transporte. A interrupção desses serviços, mesmo que breve, pode gerar impactos diretos na segurança e na qualidade de vida urbana. Nos ambientes de nuvem responsáveis por serviços urbanos críticos, falhas de infraestrutura como término inesperado de instâncias, corrupção de rede e exaustão de recursos podem se propagar rapidamente entre serviços acoplados, comprometendo a continuidade de operações essenciais à população. Diante desse cenário, a Engenharia do Caos consolidou-se como metodologia de referência para avaliação empírica de resiliência, fundamentada na injeção controlada de falhas em ambientes de produção, com o objetivo de observar o comportamento do sistema sob condições adversas e identificar fragilidades antes que estas se manifestem espontaneamente. Ferramentas amplamente adotadas pela indústria, como AWS Fault Injection Simulator, têm sido utilizadas para simular cenários realistas de falha em ambientes baseados em orquestração de contêineres, favorecendo sua aplicação também no contexto de plataformas de cidades inteligentes construídas sobre arquiteturas de nuvem. Não obstante os avanços observados, a literatura ainda aponta lacunas relevantes quanto à realização de experimentos de injeção de falhas que atravessem múltiplos provedores de nuvem, à padronização de métricas de resiliência que permitam a comparação sistemática entre plataformas distintas. Este trabalho propõe investigar, por meio de experimentação sistemática de injeção de falhas, a

resiliência de plataformas de cidades inteligentes baseadas em nuvem, com ênfase na simulação de cenários de falha representativos de condições reais de operação, e na definição de métricas quantitativas de avaliação de impacto e recuperação.

Referências:

BAGEHORN, F., RIOS, J., JHA, S., FILEPP, R., SHWARTZ, L., ABE, N., YANG, X. A fault injection platform for learning AIOps models. In: Proceedings of the 37th IEEE/ACM International Conference on Automated Software Engineering (ASE), 2023.

IKEUCHI, H., WATANABE, A., TAKAHASHI, Y. Coverage based failure injection toward efficient chaos engineering. In: ICC 2023 – IEEE International Conference on Communications, p. 4571–4577, 2023.

PALACIOS CHAVARRO, J. et al. Chaos engineering 2.0: a review of AI-driven, policy- guided resilience for multi-cloud systems. ResearchGate preprint, 2025.

PROJETO 9

Orientador: Ermeson Carneiro de Andrade

Título: Análise de Desempenho e Dependabilidade de Sistemas Computacionais e Inteligentes

Vagas: 2

Resumo:

A crescente adoção de sistemas distribuídos, sistemas ciber-físicos (Cyber-Physical Systems – CPS) e aplicações baseadas em Inteligência Artificial tem criado novos desafios relacionados ao desempenho, à dependabilidade e à operação contínua desses sistemas. O atual ecossistema de IA envolve desde modelos de aprendizado de máquina e Grandes Modelos de Linguagem (Large Language Models – LLMs) até sistemas de Geração Aumentada por Recuperação (Retrieval-Augmented Generation – RAG), agentes de IA e mecanismos de integração como o Protocolo de Contexto de Modelo (Model Context Protocol – MCP). Nesses ambientes, falhas, sobrecarga, degradação de recursos, envelhecimento de software (software aging), timeouts e mudanças na carga de trabalho podem comprometer a disponibilidade, a confiabilidade e o desempenho do serviço. Trabalhos anteriores têm explorado a modelagem de dependabilidade de CPS [1], a avaliação integrada de desempenho e disponibilidade em sistemas inteligentes com IA [2] e os efeitos do software aging em sistemas baseados em LLMs [3]. Nesse contexto, este projeto busca desenvolver métodos, modelos, técnicas e ferramentas para avaliar e melhorar o desempenho e a dependabilidade de sistemas computacionais e inteligentes, combinando, conforme o problema, experimentação, simulação, Inteligência Artificial e modelagem estocástica. As pesquisas poderão considerar métricas como disponibilidade, confiabilidade, tempo de resposta, throughput, utilização de recursos e probabilidade de conclusão bem-sucedida de tarefas, além de estratégias de redundância, recuperação, adaptação, rejuvenescimento e otimização.

[1] E. Andrade, B. Nogueira, G. Callou, and G. Alves. Dependability Analysis of a Cyber-Physical System for Smart Environments. *Concurrency and Computation: Practice and Experience*, 2019.

[2] Q. Zhang, F. Machida, and E. Andrade. SPADE: Simulator-Assisted Performability Design for UAV-Based Monitoring Systems. *Future Generation Computer Systems*, 2026.

[3] C. Santos, F. Machida, and E. Andrade. Experimental Investigation of Memory-Related Software Aging in LLM Systems. *Journal of Systems and Software*, 2026.

PROJETO 10

Orientador: Fernando Antônio Aires Lins

Título: Desenvolvimento e Avaliação de Abordagens e Soluções para Resposta a Incidentes Cibernéticos**Vagas: 3****Resumo:**

O crescimento da frequência, da complexidade e do impacto dos incidentes cibernéticos tem exigido das organizações capacidades cada vez mais estruturadas para preparação, detecção, resposta e recuperação. Entretanto, muitas instituições ainda apresentam dificuldades para avaliar objetivamente sua maturidade, identificar lacunas em seus processos e priorizar ações de melhoria. Este projeto tem como objetivo desenvolver e avaliar abordagens, métodos e soluções tecnológicas voltados ao aprimoramento da resposta a incidentes cibernéticos, considerando aspectos técnicos, organizacionais, humanos e de governança. As pesquisas desenvolvidas no âmbito do projeto poderão contemplar a proposição e a aplicação de modelos para a análise da maturidade em resposta a incidentes, incluindo indicadores, métricas, instrumentos diagnósticos e mecanismos de acompanhamento da evolução organizacional. Outra possibilidade de investigação envolve o desenvolvimento de um framework de resposta orientado por Business Impact Analysis (BIA), capaz de incorporar informações atualizadas sobre processos de negócio, ativos, dependências, criticidade e impactos, apoiando a priorização das decisões durante um incidente cibernético. O projeto também possibilita a concepção, implementação e avaliação de ferramentas e abordagens de apoio à melhoria contínua da maturidade organizacional, tais como sistemas de diagnóstico, mecanismos de recomendação, painéis de acompanhamento, exercícios de simulação, playbooks e soluções baseadas em inteligência artificial. Espera-se que os resultados contribuam para tornar a resposta a incidentes mais sistemática, mensurável e alinhada aos objetivos institucionais, produzindo conhecimento científico, artefatos tecnológicos e evidências empíricas aplicáveis a organizações públicas e privadas.

PROJETO 11**Orientador: Fernando Antônio Aires Lins****Co-orientador: Robson Wagner Albuquerque de Medeiros****Título: Avaliação da Segurança de Sistemas de Gestão Hospitalar Veterinária Apoiada por Inteligência Artificial****Vagas: 1****Resumo:**

A crescente complexidade dos sistemas de software e o aumento das ameaças cibernéticas exigem abordagens mais eficientes para identificar e mitigar vulnerabilidades. Sistemas de gestão hospitalar veterinária, por lidarem com dados sensíveis e críticos, tornam-se alvos potenciais de ataque, o que demanda soluções de segurança robustas.

Este projeto propõe o desenvolvimento e aplicação de métodos baseados em Inteligência Artificial (IA) para apoiar a avaliação da segurança desse tipo de sistema. O objetivo é investigar como modelos de IA podem auxiliar na detecção precoce de falhas, na classificação de riscos e na recomendação de ações corretivas. Espera-se que os resultados contribuam para a criação de ferramentas mais precisas e automatizadas, capazes de aumentar a resiliência e confiabilidade dessas soluções, além de oferecer uma abordagem replicável para diferentes sistemas de gestão hospitalar veterinária, fortalecendo a segurança da informação no setor.

PROJETO 12**Orientador: Fernando Antônio Aires Lins**

Co-orientador: Rafael Perazzo Barbosa Mota

Título: Segurança Ofensiva Aplicada à Análise de Causa Raiz em Incidentes Cibernéticos

Vagas: 1

Resumo:

A crescente complexidade dos ataques cibernéticos tem dificultado a identificação precisa das vulnerabilidades, falhas de configuração e deficiências de controle que possibilitam a ocorrência e a progressão dos incidentes. Embora ferramentas de monitoramento e técnicas de análise forense permitam coletar evidências e compreender parte das ações executadas pelo atacante, ainda persistem desafios relacionados à reconstrução completa do caminho de ataque e à determinação de suas causas. Nesse contexto, técnicas de segurança ofensiva podem complementar a investigação ao permitir a reprodução controlada de ações maliciosas e a validação experimental das hipóteses formuladas durante a resposta ao incidente. O projeto tem como objetivo investigar e desenvolver métodos, abordagens e ferramentas que integrem técnicas de segurança ofensiva aos processos de resposta a incidentes cibernéticos. A proposta poderá envolver atividades como reconstrução de caminhos de ataque, emulação controlada de adversários, validação de vulnerabilidades exploradas, análise de movimentação lateral, avaliação dos mecanismos de persistência e identificação das condições técnicas e organizacionais que contribuíram para o incidente. Pretende-se, assim, apoiar a análise de causa raiz e diferenciar o vetor inicial de comprometimento das fragilidades que permitiram o agravamento e a propagação do ataque.

PROJETO 13

Orientador: Filipe Rolim Cordeiro

Título: Equidade e Benefício Clínico em IA para Imagens Médicas

Vagas: 1

Resumo:

Modelos de aprendizado profundo já conseguem detectar doenças como o câncer de pele e de mama, com desempenho comparável ao de especialistas. No entanto, um problema importante permanece: esses modelos podem funcionar muito bem para alguns grupos de pacientes e significativamente pior para outros. Diferenças relacionadas a características como idade, sexo e raça podem levar justamente os pacientes mais vulneráveis a receber diagnósticos menos confiáveis. Grande parte dos trabalhos atuais avalia a equidade por meio de métricas estatísticas de desempenho, como AUC, sensibilidade e especificidade. Essas métricas, porém, nem sempre refletem o impacto real das decisões clínicas. Na prática, é preciso considerar fatores como a prevalência da doença, o limiar utilizado para recomendar uma biópsia e, principalmente, o fato de que não detectar um câncer pode ter consequências muito mais graves do que realizar um exame adicional desnecessário. Este projeto de mestrado pretende investigar a equidade em modelos de visão computacional sob a perspectiva dos benefícios clínicos para diferentes grupos de pacientes. Para isso, utilizaremos os conceitos de Decision Curve Analysis e Net Benefit, aproximando a avaliação da fairness das decisões que efetivamente ocorrem na prática médica.

PROJETO 14

Orientador: Filipe Rolim Cordeiro

Título: Aprendizado de Máquina Robusto a Anotações Ruidosas em Imagens Médicas

Vagas: 1

Resumo:

Sistemas de inteligência artificial já conseguem alcançar desempenho próximo ao de especialistas em diversas tarefas de diagnóstico por imagem. Para aprender, entretanto, esses modelos dependem de milhares de exames previamente anotados por especialistas, cujas anotações nem sempre estão corretas. Casos ambíguos, doenças raras e diferenças de interpretação podem levar especialistas a discordar sobre o diagnóstico de uma mesma imagem. Surge, então, uma pergunta importante: o que acontece quando treinamos uma rede neural com anotações que contêm erros? Existe uma extensa área de pesquisa dedicada a esse problema, conhecida como Learning with Noisy Labels (LNL). Entretanto, há uma diferença importante entre a forma como o problema costuma ser estudado e o que ocorre na prática médica. Grande parte dos trabalhos considera taxas elevadas de ruído artificial, frequentemente introduzidas aleatoriamente em bases de dados grandes e relativamente balanceadas. No cenário médico, os erros costumam ser menos frequentes, concentrando-se nos casos mais difíceis e em bases pequenas e fortemente desbalanceadas. Nessas condições, uma estratégia criada para identificar e remover exemplos potencialmente incorretos pode acabar eliminando casos difíceis ou raros que estavam corretamente anotados e são importantes para o diagnóstico. Este projeto de mestrado pretende investigar estratégias para lidar com o aprendizado a partir de anotações ruidosas em cenários mais próximos aos encontrados em aplicações médicas reais, combinando ruído de anotação, dificuldade das imagens, desbalanceamento entre classes e escassez de dados.

PROJETO 15

Orientador: Filipe Rolim Cordeiro

Título: Generalização de Modelos de Visão Computacional em Imagens de Ultrassom de Mama

Vagas: 1

Resumo:

O câncer de mama é o mais incidente entre mulheres no Brasil, e o ultrassom é a modalidade de imagem mais acessível para o diagnóstico, inclusive no SUS. Nos últimos anos, redes neurais profundas atingiram valores de acurácia de 97–99% na classificação de lesões mamárias em imagens de ultrassom, quando analisadas em imagens de um mesmo hospital. No entanto, uma mudança de hospital significa mudanças no equipamento de ultrassom, no protocolo de aquisição, na população de pacientes e nas características das imagens. Como consequência, um modelo que apresenta excelente desempenho nas imagens utilizadas durante seu desenvolvimento pode sofrer uma queda significativa quando aplicado a dados de uma instituição diferente. Este projeto analisará arquiteturas modernas de Deep Learning aplicadas a diferentes bases públicas de ultrassom de mama. Em vez da avaliação tradicional, em que treinamento e teste utilizam imagens da mesma base, investigaremos cenários cross-dataset e cross-device, nos quais o modelo será testado em imagens provenientes de hospitais ou de equipamentos que não foram utilizados durante o treinamento. A partir dos problemas identificados, serão investigadas estratégias que reduzam a dependência de características específicas de cada domínio e incentivem o aprendizado de informações mais relacionadas à própria lesão, como sua forma e estrutura. O objetivo do trabalho é desenvolver modelos que mantenham bom desempenho mesmo quando avaliados em hospitais e equipamentos nunca vistos durante o treinamento.

PROJETO 16

Orientador: Gabriel Alves de Albuquerque Júnior

Título: Trajetória Estudantil Justa e Gestão Acadêmica Ética com Small Language Models e Grafos de Conhecimento

Vagas: 2

Resumo:

Como garantir que um sistema de apoio à vida acadêmica seja ao mesmo tempo justo para todos os estudantes e útil para quem toma decisões institucionais? Essa é a questão central deste projeto, que parte da estrutura curricular da UFRPE para construir respostas concretas. O ponto de partida é a representação do currículo institucional como um knowledge graph, conectando cursos, componentes, competências, pré-requisitos e trajetórias acadêmicas em uma estrutura que torna explícitas relações que permanecem invisíveis em documentos convencionais. Sobre esse grafo, técnicas de Knowledge Tracing permitem modelar a evolução do domínio estudantil ao longo do tempo, identificando lacunas, estimando trajetórias e sinalizando riscos de forma individualizada. A abordagem GraphRAG transforma esse conhecimento estruturado em base para um assistente institucional capaz de responder perguntas complexas navegando as relações do grafo, algo que métodos convencionais de recuperação não conseguem fazer. Para que essa solução seja implantável em um ambiente universitário com restrições reais de infraestrutura e privacidade, o projeto adota Small Language Models ajustados ao domínio acadêmico da instituição, em substituição a modelos de grande porte que dependem de serviços externos. A dimensão ética atravessa todo o projeto. Os modelos de rastreamento de trajetória e os preditores de desempenho e evasão do sistema SABIA da UFRPE serão avaliados e aprimorados com métricas de fairness, monitorando e mitigando ativamente vieses associados a gênero, renda e raça, para que as recomendações produzidas não reproduzam as desigualdades presentes nos dados históricos.

PROJETO 17

Orientador: Gabriel Alves de Albuquerque Júnior

Título: Expressões Culturais do Semiárido Brasileiro: Mapeamento e Recuperação com Knowledge Graphs e GraphRAG

Vagas: 1

Resumo:

O semiárido brasileiro é uma das regiões culturalmente mais ricas e singulares do país. Suas expressões, como a poesia, o cordel, o repente, o música, a dança, o artesanato em cerâmica, couro e fibras, além de festas populares e saberes tradicionais transmitidos oralmente, formam um tecido identitário complexo, marcado por territorialidades, histórias e sentidos que resistem a qualquer simplificação. Apesar dessa riqueza, essas manifestações permanecem fragmentadas em fontes dispersas e raramente representadas de forma estruturada, o que limita seu estudo, sua preservação e sua difusão em ambientes digitais. Este projeto propõe superar essa fragmentação por meio da construção de um knowledge graph que capture entidades como artistas, grupos, municípios, eventos, práticas e artefatos, bem como as relações semânticas, históricas e territoriais entre elas. O grafo será construído a partir de metodologias de extração de conhecimento aplicadas a fontes heterogêneas, incluindo acervos digitais, textos jornalísticos, registros institucionais e depoimentos transcritos. Essa estrutura relacional permite rastrear como manifestações culturais se difundem, se influenciam mutuamente e se transformam ao longo do tempo e do território, revelando conexões que uma coleção de documentos isolados jamais tornaria visível. Sobre esse grafo, serão desenvolvidas e avaliadas estratégias de recuperação baseadas em GraphRAG, habilitando consultas em linguagem natural que exploram simultaneamente a semântica dos textos e as relações estruturais do

grafo para responder perguntas que nenhuma busca convencional alcança. O projeto contribui para a preservação do patrimônio imaterial regional, para o avanço metodológico na aplicação de knowledge graphs e GraphRAG em domínios culturais com baixa representação computacional e para o desenvolvimento de sistemas de acesso a esse conhecimento por pesquisadores, educadores e comunidades.

PROJETO 18

Orientador: George Augusto Valença Santos

Título: Proteção de Crianças e Adolescentes em Plataformas e Tecnologias Digitais.

Vagas: 4

Resumo:

Crianças e adolescentes utilizam plataformas e tecnologias digitais para aprender, socializar, criar, jogar e acessar diferentes formas de entretenimento. Ao mesmo tempo, esses ambientes podem expô-los a riscos relacionados à privacidade, segurança, bem-estar, autonomia e desenvolvimento, incluindo práticas de design manipulativo, exposição a conteúdos inadequados, interações potencialmente danosas, exploração comercial e usos inadequados de dados pessoais. Esta linha de pesquisa busca investigar, projetar e avaliar tecnologias digitais sob a perspectiva da proteção e dos direitos de crianças e adolescentes. Os trabalhos podem envolver diferentes tipos de sistemas e serviços, como redes sociais, jogos digitais, plataformas educacionais, sistemas baseados em Inteligência Artificial e agentes conversacionais, brinquedos conectados e outras tecnologias emergentes. As pesquisas poderão abordar temas como design manipulativo e persuasivo, privacidade e proteção de dados, verificação e garantia de idade, controles parentais, segurança por design, inteligência artificial, riscos algorítmicos e mecanismos de proteção e bem-estar digital. Também poderão investigar como características de interface, modelos de negócio e decisões de design influenciam as experiências e os comportamentos de crianças e adolescentes. Metodologicamente, os trabalhos poderão combinar abordagens da Engenharia de Software e Interação Humano-Computador, incluindo estudos empíricos com plataformas, análise de interfaces e funcionalidades, experimentos, estudos com usuários, desenvolvimento e avaliação de protótipos, revisão da literatura e análise de requisitos e regulamentações. Busca-se produzir evidências e soluções que contribuam para o desenvolvimento de tecnologias mais seguras, responsáveis e adequadas aos direitos e às necessidades de crianças e adolescentes.

PROJETO 19

Orientador: Gilberto Amado de Azevedo Cysneiros Filho

Título: Estratégias de Desenvolvimento de Jogos e Aplicações Educacionais usando Abordagens Gamificadas Integrando Engenharia de Software, UX e Neurociência.

Vagas: 2

Resumo:

O uso de tecnologias aplicadas na educação formal e não formal vem aumentando e a grande questão é como integrar o aprendizado e ludicidade através de abordagens que respeitem o usuário. Nossa linha de pesquisa é investigar e propor novas estratégias, diretrizes e metodologias para que o desenvolvimento desses softwares respeite os indivíduos e a sociedade tanto do aspecto da LGPD como dos Objetivos do Desenvolvimento Sustentável. Em particular, nosso escopo é explorar novas tecnologias como IA Generativa, Realidade Virtual e Aumentada e principalmente como os elementos de gamificação

combinados com design da interface alinhados ao conhecimento da neurociência ajudam a criar softwares eficazes, mas que respeitem a saúde mental dos indivíduos e atendam as restrições da LGPD (e outras regulamentações de domínio específico) e também sejam guiados pelos objetivos de desenvolvimento sustentável. Nosso entendimento é que a tecnologia vem como uma ferramenta para auxiliar o aprendizado e a evolução da sociedade, mas se usado sem uma metodologia de desenvolvimento científico pode também trazer efeitos negativos na sociedade, mesmo quando o objetivo é ser uma solução de melhoria. Desta forma, buscamos também respeitar o conhecimento dos domínios específicos usando técnicas de Engenharia de Requisitos para exemplificar temos explorar diversas áreas da educação como: Letramento, História e Cultura, Biologia e Pensamento Computacional. Em relação à neurociência, nós investigamos como os elementos de design visuais, sonoros, animações, interações digitais junto com o desenvolvimento cognitivo, uso de gamificação e estratégias educacionais influenciam o aprendizado em geral.

PROJETO 20

Orientador: Gustavo Rau de Almeida Callou

Título: Uma Abordagem integrada de IoT, Cloud Computing e IA para Mitigar os Problemas da Segurança Pública

Vagas: 1

Resumo:

A segurança pública é fundamental para o funcionamento saudável e eficiente de uma sociedade, abrangendo um conjunto de medidas e ações destinadas a proteger a ordem, a integridade física e os direitos dos cidadãos. A segurança pública não se limita apenas à prevenção e combate à criminalidade, mas também envolve a promoção do bem-estar social e o fortalecimento da confiança na comunidade. Algumas das razões pelas quais a segurança pública é crucial se encontra relacionada com a proteção dos cidadãos, uma vez que a principal função da segurança pública é garantir a segurança e proteção dos cidadãos. Isso inclui a prevenção e a repressão de atividades criminosas que possam ameaçar a vida, a propriedade e a paz das pessoas. É importante destacar que a segurança pode ter uma relação direta com o desenvolvimento econômico, onde a presença de uma segurança pública robusta cria um ambiente propício para o desenvolvimento econômico. Empresas e investidores são mais propensos a se estabelecer em áreas seguras. O presente trabalho tem como objetivo a proposição de uma abordagem que irá fazer uso de dispositivos de internet das coisas (IoT) em conjunto com computação em nuvem com a finalidade de proporcionar um monitoramento automático e em tempo real de regiões críticas do ponto de vista da segurança pública. A estratégia proposta irá auxiliar o efetivo policial a identificar possíveis ocorrências e, assim, às autoridades já vão poder ter ciência do ocorrido antes mesmo da população reportar à polícia ou entrar em contato com serviços de emergência. Dessa forma, espera-se conseguir reduzir e melhor posicionar o efetivo policial para os locais com maior criminalidade ocorrendo. Para que isso seja possível, esse projeto irá fazer uso de dispositivos de IoT (ex., placas Raspberry Pi ou Orange Pi) como computação de borda (edge computing) que poderão ser instalados em viaturas policiais, praças e parques públicos, transporte público, postes em vias públicas nos bairros com maior incidência de crimes. Além disso, iremos incorporar drones com capacidade de aquisição de imagens ao sistema. Essas placas vão ser gerenciadas por uma nuvem privada e vão ficar recebendo em tempo real as imagens de diversas câmeras. O diferencial desse trabalho é que tanto câmeras fixas como moveis vão ser utilizadas para auxiliar na detecção de possíveis crimes. Essa detecção terá o auxílio de inteligência artificial (IA) e com o suporte em tempo real provido pela edge computing será possível fazer com que o efetivo policial se desloque mais rápido à

ocorrência, auxiliando na solução efetiva da criminalidade. Além disso, para garantir que as imagens estejam sempre disponíveis, esse projeto também irá contar com técnicas de tolerância a falhas a fim de maximizar a disponibilidade e reduzir o tempo de resposta demandado. Por exemplo, redundância com multi edge e na cloud serão utilizadas. Particularmente, a redundância na edge representa um desafio que será estudado e solucionado nesse trabalho no contexto em que uma câmera móvel pode estar mais próxima de uma edge em determinado momento, mas ao se deslocar, ela pode ter que mudar para outra edge. Para isso, modelos em redes de Petri estocásticas também serão propostos para auxiliar na quantificação de tais métricas de desempenho e dependabilidade.

Referências:

- [1] SILVA, Carolina Vanessa Meireles. Inteligência policial e os desafios da segurança pública nas cidades 4.0. 2025.
- [2] MAHOR, Vinod et al. IoT and artificial intelligence techniques for public safety and security. In: Smart urban computing applications. River Publishers, 2023. p. 111-126.

PROJETO 21

Orientador: Gustavo Rau de Almeida Callou

Título: Algoritmos Inteligentes e Baseados em Grafos Aplicados à Investigação Financeira

Vagas: 1

Resumo:

A lavagem de dinheiro representa uma ameaça crítica à integridade dos sistemas financeiros e está diretamente associada à sustentação de diversas atividades criminosas, como tráfico de drogas, corrupção e financiamento ao terrorismo [4][6]. Diante da complexidade e do grande volume de dados envolvidos, o enfrentamento eficaz desse crime exige soluções tecnológicas avançadas e automatizadas. Neste contexto, este projeto propõe o desenvolvimento de uma abordagem baseada na modelagem de grafos de fluxo, redes neurais e algoritmos de inteligência computacional, com foco na detecção de indícios de lavagem de dinheiro. Os grafos de fluxo permitem representar de forma estruturada as relações entre entidades financeiras, evidenciando padrões e anomalias em movimentações suspeitas [1][2][3]. Integrados a técnicas de inteligência artificial, esses grafos possibilitam análises automatizadas e a identificação precoce de comportamentos atípicos, muitas vezes imperceptíveis por métodos tradicionais [5]. O sistema proposto buscará automatizar etapas investigativas, oferecendo aos analistas uma ferramenta visual e analítica robusta para cruzamento e interpretação de dados financeiros. O projeto está inserido na 6ª rodada do Programa Cientista Arretado da FACEPE (Edital 14/2024), em colaboração com a Secretaria de Defesa Social de Pernambuco (SDS-PE), atendendo a uma demanda prática do Laboratório de Combate à Lavagem de Dinheiro da Diretoria de Inteligência da Polícia Civil de Pernambuco (DINTEL-PCPE). A proposta está alinhada às diretrizes da Carta Circular nº 4.001/2020 do Banco Central do Brasil [7], garantindo aderência às normativas legais e operacionais. Espera-se, assim, contribuir para o aprimoramento da eficiência, precisão e celeridade das investigações financeiras, promovendo o uso estratégico de ciência de dados no combate à criminalidade econômica no Brasil.

Referências:

- [1] Blanuša, Jovan, et al. Graph Feature Preprocessor: Real-time Subgraph-based Feature Extraction for Financial Crime Detection. Proceedings of the 5th ACM International Conference on AI in Finance. 2024.
- [2] Wójcik, Filip. An Analysis of Novel Money Laundering Data Using Heterogeneous Graph Isomorphism Networks. FinCEN Files Case Study. Econometrics. Ekonometria. Advances in Applied Data Analytics. 2024.

- [3] WANG, Shaojiang et al. Structural entropy minimization combining graph representation for money laundering identification. *International Journal of Machine Learning and Cybernetics*, v. 15, n. 9, p. 3951-3968, 2024.
- [4] Da Cunha, Bruno Requião. Neutralização seletiva de alvos topológicos de alto retorno em facções criminosas. *Revista Brasileira de Ciências Policiais*, v. 12, n. 4, p. 53-73, 2021.
- [5] Morselli, Carlo. *Inside criminal networks*. New York: Springer, 2009.
- [6] Da Cunha, Bruno Requião; GONÇALVES, Sebastián. Topology, robustness, and structural controllability of the Brazilian Federal Police criminal intelligence network. *Applied network science*, v. 3, p. 1-20, 2018.
- [7] BANCO CENTRAL DO BRASIL. Carta Circular nº 4.001 de 29 de janeiro de 2020. Disponível em: https://normativos.bcb.gov.br/Lists/Normativos/Attachments/50911/C_Circ_4001_v2_P.pdf

PROJETO 22

Orientador: Gustavo Rau de Almeida Callou

Co-Orientador: Thiago Valentim Bezerra

Título: Classificação de Subtipos de Câncer a partir de Dados de Expressão Gênica Utilizando Técnicas de Aprendizado de Máquina

Vagas: 1

Resumo:

O câncer é uma das principais causas de morbimortalidade em todo o mundo, e a variação nos programas transcricionais das células tumorais é responsável por grande parte da diversidade biológica observada entre os tumores [1]. Diferentes subtipos moleculares de um mesmo tipo de câncer podem apresentar prognósticos e respostas terapêuticas distintas mesmo quando os tumores são histologicamente semelhantes, como demonstrado em estudos que associaram padrões de expressão gênica de carcinomas de mama a subclasses tumorais com implicações clínicas diretas [2]. O sequenciamento de RNA (RNA-Seq) permitiu mapear e quantificar transcriptomas completos com maior resolução e sensibilidade do que métodos anteriores, gerando, em contrapartida, um grande volume de dados de alta dimensionalidade e sujeitos a ruído [3]. Este projeto propõe uma abordagem baseada em aprendizado de máquina e redes neurais para classificar subtipos de câncer a partir de perfis de expressão gênica. Os dados de expressão gênica, obtidos por sequenciamento de RNA, descrevem de forma quantitativa os padrões transcricionais de cada amostra, permitindo distinguir subtipos tumorais a partir de assinaturas moleculares específicas [1-2]. A viabilidade de algoritmos de aprendizado de máquina para essa tarefa foi demonstrada já no final da década de 1990, quando técnicas de descoberta e predição de classes aplicadas a dados de expressão gênica possibilitaram classificar tipos de câncer de forma automatizada, sem depender exclusivamente de conhecimento biológico prévio sobre os genes envolvidos [4]. Esse tipo de abordagem, hoje estendida a arquiteturas de redes neurais mais complexas, permite identificar relações entre genes que não seriam facilmente percebidas por métodos estatísticos tradicionais.

Referências:

- [1] PEROU, C. M. et al. Molecular portraits of human breast tumours. *Nature*, v. 406, p. 747–752, 2000.
- [2] SØRLIE, T. et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *PNAS*, v. 98, n. 19, p. 10869–10874, 2001.
- [3] MORTAZAVI, A. et al. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods*, v. 5, p. 621–628, 2008.

[4] GOLUB, T. R. et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science, v. 286, n. 5439, p. 531–537, 1999.

PROJETO 23

Orientador: João Paulo Silva do Monte Lima

Título: Classificação do peso corporal de frangas de postura utilizando imagens de profundidade e RGB associado a aprendizado de máquina

Vagas: 1

Resumo:

O acompanhamento do peso corporal de aves de postura é essencial para assegurar capacidade de produção de ovos. Porém, os métodos de monitoramento manuais têm sido negligenciados em função da necessidade de manuseio e pesagem individual das aves em grandes lotes, sendo necessárias alternativas para automatizar e facilitar esse processo. A combinação de imagens, visão computacional e aprendizado de máquina é uma ferramenta útil que permite a predição de peso de animais de produção, entretanto, a classificação de aves de postura de acordo com o peso corporal ainda é inexplorada, apesar de necessária. O objetivo desse projeto será classificar frangas de postura de acordo com a categoria de peso corporal utilizando imagens de profundidade e RGB associado a técnicas de visão computacional e aprendizado de máquina. Nesse sentido, hipotetizou-se que será possível classificar frangas poedeiras em categorias de peso corporal utilizando imagens, visão computacional e aprendizado de máquina obtendo boas métricas de precisão e acurácia. Serão utilizadas 200 aves de postura com 1 dia de idade. Às 6 semanas de idade, 45 aves serão selecionadas, pesadas e classificadas em três categorias de peso corporal (leves, médias e pesadas). Cada ave será submetida à captura de imagens por meio de filmagem de 10 segundos com câmera de profundidade e RGB em ângulo dorsal e lateral. As filmagens serão processadas por meio de algoritmo de extração de frames, para a aquisição de 500 a 1500 capturas. As imagens serão processadas utilizando técnicas de segmentação, ajuste paramétrico e remoção de ruído e os pixels das capturas serão normalizados. Após o processamento, as imagens serão submetidas ao aprendizado de máquina supervisionado, seguindo o modelo composto por coleta dos dados, pré-processamento, construção do modelo e avaliação do modelo. Os modelos testados serão avaliados segundo as métricas de acurácia, precisão, recall, escore F-1, AUC e coeficiente de correlação de Matthews.

PROJETO 24

Orientador: Lucas Albertins de Lima

Título: Dos modelos a sistemas executáveis: Construindo uma máquina virtual para SysML v2

Vagas: 2

Resumo:

Como transformar um modelo de engenharia em uma representação que não apenas descreva um sistema, mas que possa ser executada, observada, verificada e explorada antes mesmo de sua implementação? Essa é uma das questões centrais para a evolução da Engenharia de Sistemas Baseada em Modelos (MBSE). A linguagem SysML v2 representa um importante avanço nesse contexto, ao oferecer uma semântica mais rica e precisa para a descrição de sistemas e de seu comportamento. Entretanto, ainda existem desafios fundamentais para transformar esses modelos em artefatos efetivamente executáveis, permitindo analisar seu comportamento, explorar diferentes alternativas de projeto e verificar suas propriedades de forma integrada. Este projeto investiga uma nova abordagem para superar essas limitações: uma máquina virtual

para SysML v2 capaz de executar e analisar modelos diretamente, preservando sua semântica e permitindo sua exploração dinâmica. A proposta, denominada Living SysML v2 Blueprint, combina três capacidades fundamentais: execução do comportamento do sistema, observação detalhada das execuções e verificação de propriedades do modelo. Por meio da API de Execução, Observação e Controle Transparente (TEOC), pretende-se possibilitar a inspeção da semântica do modelo e a interoperabilidade com diferentes ferramentas sem depender de transformações que possam comprometer a fidelidade do modelo original. Um aspecto particularmente inovador é a possibilidade de explorar um “multiverso de projeto”, no qual diferentes decisões e configurações podem ser experimentadas e comparadas de maneira sistemática. Essa capacidade abre novas possibilidades para a identificação antecipada de problemas, análise de alternativas, explicabilidade das decisões de engenharia e, futuramente, integração com técnicas de inteligência artificial para apoiar o processo de projeto. O aluno de mestrado poderá atuar em desafios de semântica de linguagens, execução de modelos, verificação formal, exploração de estados e desenvolvimento de software, contribuindo para a construção e validação de uma plataforma experimental de nova geração para MBSE. O projeto será desenvolvido em colaboração com um grupo de pesquisa do Instituto ENSTA Bretagne, na França, proporcionando uma oportunidade de pesquisa internacional em uma área emergente da Engenharia de Sistemas. A equipe brasileira contribuirá diretamente para o desenvolvimento da máquina virtual e de sua semântica executável para SysML v2, além de atuar na definição e implementação de mecanismos de verificação e validação da plataforma.

PROJETO 25

Orientador: Lucas Albertins de Lima

Título: Dos Dados ao Conhecimento: Ontologias, Grafos de Conhecimento e Inteligência Artificial para a Construção de Sistemas

Vagas: 2

Resumo:

Como transformar grandes volumes de dados dispersos em conhecimento útil para apoiar o desenvolvimento de sistemas complexos? Em muitos projetos de sistemas, informações relevantes encontram-se distribuídas entre documentos, bancos de dados, modelos, sistemas legados e diferentes especialistas. Além disso, parte importante do conhecimento necessário para compreender um sistema permanece implícita ou está na experiência de profissionais, tornando sua captura, integração e reutilização um desafio significativo. A Engenharia de Conhecimento oferece fundamentos para enfrentar esse problema por meio de técnicas capazes de capturar, representar, organizar, integrar e disponibilizar conhecimento de forma estruturada. Ontologias, grafos de conhecimento e mecanismos de raciocínio permitem conectar informações provenientes de diferentes fontes e criar uma visão integrada do domínio. Mais recentemente, o avanço dos Large Language Models (LLMs) e de agentes inteligentes abre novas possibilidades para automatizar etapas desse processo, como a extração e normalização de informações, a construção de conhecimento e a interação com bases de conhecimento. Este projeto investiga como combinar Engenharia de Conhecimento e Inteligência Artificial para transformar dados heterogêneos e conhecimento especializado em recursos computacionais capazes de apoiar o desenvolvimento e a análise de sistemas. Entre as possibilidades de investigação estão a construção e evolução de ontologias, a representação de conhecimento por meio de grafos, a integração de fontes heterogêneas utilizando grafos federados e o uso de LLMs para extração, normalização e enriquecimento de conhecimento. O projeto oferece diferentes desafios de pesquisa, permitindo explorar questões como: Como representar

conhecimento de forma que máquinas possam interpretá-lo e raciocinar sobre ele? Como integrar informações provenientes de fontes heterogêneas sem perder seu significado? Como utilizar LLMs para extrair conhecimento de documentos e dados não estruturados? Como combinar conhecimento estruturado e modelos de IA para produzir respostas mais confiáveis, rastreáveis e explicáveis? O aluno de mestrado poderá atuar em um ou mais desses desafios, combinando Engenharia de Conhecimento, Inteligência Artificial, Ontologias, Knowledge Graphs, LLMs e integração de dados. Dependendo do recorte escolhido, o trabalho poderá envolver desde a concepção de modelos de conhecimento e algoritmos de raciocínio até o desenvolvimento e avaliação de protótipos baseados em IA.

O objetivo é contribuir para uma nova geração de sistemas capazes não apenas de armazenar dados, mas de compreender relações, integrar informações, produzir conhecimento e apoiar decisões de forma inteligente e explicável.

PROJETO 26

Orientador: Marcelo Luiz Monteiro Marinho

Título: Segurança Psicológica em Equipes Ágeis de Desenvolvimento de Software: uma Investigação Baseada em Emoções e Análise de Sentimentos

Vagas: 1

Resumo:

No desenvolvimento ágil de software, a segurança psicológica é fundamental para a comunicação aberta, a colaboração e a disposição dos membros da equipe para expressar ideias, dúvidas e erros sem receio de consequências negativas [1]. Estudos recentes apontam sua relação com comportamentos importantes para a qualidade do software, como a aprendizagem com erros, a colaboração e a resolução coletiva de problemas [2]. Apesar de sua relevância, ainda existem desafios para compreender e acompanhar a forma como a segurança psicológica se manifesta e se modifica no cotidiano de trabalho das equipes [3].

As emoções e os sentimentos expressos pelos membros das equipes podem oferecer uma perspectiva complementar a essa análise, uma vez que estão relacionados à comunicação, à colaboração e à própria segurança psicológica [4]. Além disso, situações recorrentes no desenvolvimento ágil, como conflitos, problemas na atribuição de papéis e expectativas organizacionais, podem afetar a segurança psicológica e gerar consequências, tais como desmotivação e redução da produtividade [5]. Este projeto de mestrado propõe investigar como a análise de sentimentos e emoções, apoiada por técnicas de Inteligência Artificial, pode contribuir para a identificação e o acompanhamento da segurança psicológica em equipes ágeis de desenvolvimento de software. A pesquisa buscará compreender quais sinais emocionais e padrões de sentimento podem estar associados a variações na segurança psicológica e em que medida esses sinais podem complementar instrumentos tradicionais de avaliação.

Referências:

- [1] Alami, A., Zahedi, M., & Krancher, O. (2023). Antecedents of psychological safety in agile software development teams. *Information and Software Technology*, 162, 107267.
- [2] Santana, B. et al. (2025). Psychological safety in software workplaces: A systematic literature review. *Information and Software Technology*, 187, 107838.
- [3] Marinho, M., Albuquerque, F., Amorim, L., & Sampaio, S. (2025, September). Explorando a Influência da Segurança Psicológica na Qualidade do Software em Equipes Ágeis. In *Simpósio Brasileiro de Engenharia de Software (SBES)* (pp. 60-70). SBC.

[4] Melo, A. M. B., Marinho, M., & Sampaio, S. (2025). AgileMood: A Tool for Ongoing Measurement of Psychological Safety in Agile Teams. O trabalho estabelece uma base particularmente importante para esta proposta ao integrar segurança psicológica, emoções e acompanhamento longitudinal.

[5] Monte, L., Santana, B., de Figueiredo Carneiro, G., Mendonça, M., Freire, S., & Santos, J. A. M. (2025, September). Understanding Psychological Safety in Agile Software Development Teams. In Simpósio Brasileiro de Engenharia de Software (SBES) (pp. 416-426). SBC.

PROJETO 27

Orientador: Marcelo Luiz Monteiro Marinho

Título: O Impacto da IA Generativa na Carga Cognitiva em Times Ágeis de Desenvolvimento de Software

Vagas: 1

Resumo:

A Inteligência Artificial Generativa vem sendo incorporada a diferentes atividades realizadas por times ágeis de desenvolvimento de software, apoiando tarefas como geração de código, testes, documentação, revisão e planejamento. Embora seu uso seja frequentemente associado à redução de esforço cognitivo e à liberação dos profissionais para atividades mais complexas [1,2], estudos recentes indicam que seus efeitos sobre a carga cognitiva podem ser mais complexos. Em contextos ágeis, a automação pode reduzir a carga de determinadas tarefas, ao mesmo tempo em que introduz novas demandas associadas à interação com as ferramentas, à revisão e à avaliação dos resultados produzidos pela IA [3,4]. Este projeto de mestrado propõe investigar como o uso de IA generativa influencia a carga cognitiva dos profissionais que atuam em times ágeis de desenvolvimento de software. A pesquisa buscará compreender em quais situações a IA contribui para reduzir, aumentar ou redistribuir a carga cognitiva ao longo das atividades e práticas ágeis, considerando aspectos como a revisão de artefatos gerados, a verificação de resultados, a tomada de decisão e a interação humano-IA. Pretende-se, assim, ampliar a compreensão das consequências humanas da adoção dessas tecnologias, indo além dos indicadores tradicionais de produtividade e contribuindo para formas mais adequadas de integração da IA ao trabalho cotidiano dos times ágeis.

Referências:

[1] Stray, V., Moe, N. B., Ganeshan, N., & Kobbenes, S. (2025, January). Generative ai and developer workflows: How github copilot and chatgpt influence solo and pair programming. In Proceedings of the 58th Hawaii International Conference on System Sciences.

[2] Barbala, A., Ulfsnes, R., Wivestad, V., & Stray, V. (2025). Generative AI in the workplace: Affective affordances and employee flourishing.

[3] Palla, D., Krejcar, O., Selamat, A., & Trillo, J. R. (2026, May). Cognitive Load as a Lens for Understanding AI-Augmented Workflows in Agile Teams. In 2026 IEEE Conference on Artificial Intelligence (CAI) (pp. 1653-1658). IEEE.

[4] Fujs, D., Kochovski, P., Stankovski, V., & Vavpotič, D. (2026). Key AI features to support scrum software engineering: practitioners' perspective. Empirical Software Engineering, 31(5), 138.

PROJETO 28

Orientador: Rafael Ferreira Leite de Mello

Título: Projeto e Avaliação de um Tutor Socrático Multiagente Baseado em LLM para Reduzir a Entrega Direta de Soluções em Disciplinas de Introdução à Programação

Vagas: 1

Resumo:

O uso de Grandes Modelos de Linguagem (LLMs) em disciplinas de Introdução à Programação pode auxiliar os estudantes, mas também favorecer a obtenção de códigos prontos e reduzir sua participação na construção das soluções. Este trabalho tem como objetivo projetar, implementar e avaliar um tutor socrático multiagente baseado em LLM, capaz de oferecer apoio pedagógico sem entregar diretamente soluções completas. A arquitetura proposta separa as responsabilidades de mediação pedagógica, análise técnica interna, controle de exposição e avaliação das respostas entre quatro agentes, apoiados por um modelo mínimo do estudante e por uma política de exposição em níveis progressivos. Trata-se de uma pesquisa aplicada, predominantemente qualitativa e descritiva, conduzida por meio da Design Science Research. O artefato será avaliado em cenários controlados de programação introdutória, envolvendo pedidos diretos de solução, insistência, tentativas de contornar as regras, erros de código e dificuldades para iniciar ou continuar uma atividade. As respostas serão analisadas quanto à entrega direta de soluções, à mediação socrática, à progressão das dicas e à utilidade pedagógica. Espera-se produzir um protótipo funcional e auditável que apoie o raciocínio do estudante e preserve seu papel ativo na resolução de problemas de programação.

PROJETO 29

Orientador: Rafael Ferreira Leite de Mello

Título: Construção de um Banco Público de Questões de Matemática Etiquetado por Habilidades da BNCC para Treinamento e Avaliação de Modelos de Inteligência Artificial

Vagas: 1

Resumo:

A inteligência artificial aplicada à educação exige dados de treinamento alinhados ao currículo oficial para garantir que sistemas de tutoria adaptativa, geração e correção automática de questões operem em conformidade com as diretrizes educacionais nacionais. Embora existam bancos públicos de questões de Matemática e representações abertas da Base Nacional Comum Curricular (BNCC), há uma lacuna de dados unificados e auditáveis que associem de forma confiável cada questão às suas respectivas habilidades curriculares. O objetivo deste trabalho é construir, validar e documentar um banco público de questões de Matemática da Educação Básica etiquetado por habilidades da BNCC para servir como dataset de treinamento e avaliação (fine-tuning e benchmarking) de modelos de IA educacionais. Trata-se de uma pesquisa aplicada, exploratória e de abordagem mista, composta pelo levantamento de fontes públicas, desenvolvimento de um pipeline reprodutível de etiquetagem (combinando busca por palavras-chave, classificação semântica e revisão humana), validação especialista por amostragem e condução de um experimento-piloto com um modelo de IA. Espera-se disponibilizar um dataset aberto e documentado, um pipeline de código reutilizável, um relatório de validação da qualidade dos rótulos e um artigo científico para a área de Informática na Educação. Conclui-se que o banco de dados e sua metodologia oferecem uma base sólida e auditável para o desenvolvimento responsável de tecnologias educacionais alinhadas à BNCC.

PROJETO 30

Orientador: Rafael Ferreira Leite de Mello

Título: Graph-Driven Exam Gen: Plataforma GraphRAG para Geração de Itens com Ancoragem na Matriz de Referência do ENEM

Vagas: 1

Resumo:

A geração automática de itens de avaliação por Grandes Modelos de Linguagem (LLMs) tem avançado em qualidade linguística, mas frequentemente sem garantias de alinhamento às matrizes curriculares normativas. No contexto do ENEM, com mais de 4 milhões de participantes anuais, a construção de bancos robustos de itens para Testes Adaptativos Computadorizados (CAT) exige rastreabilidade estrutural entre questão, habilidade e competência, requisito que sistemas convencionais de geração não oferecem. O objetivo deste trabalho é projetar e desenvolver uma plataforma GraphRAG que modela a Matriz de Referência completa do ENEM como grafo de propriedades, cobrindo as quatro áreas (Linguagens, Matemática, Ciências da Natureza e Ciências Humanas), suas competências e 120 habilidades, injetando esse contexto curricular estruturado no pipeline de geração de LLMs, respondendo ao problema de como garantir conformidade pedagógica em escala integral sem comprometer a autonomia do avaliador. Trata-se de pesquisa aplicada, quasi-experimental e mista, com prototipação de um pipeline de geração em larga escala, validação estrutural automatizada e avaliação por juízes especialistas conforme protocolo blinded. Métricas de coerência lexical, especificidade temática e conformidade estrutural são implementadas; avaliação pedagógica por especialistas em elaboração ENEM está planejada para fase seguinte. Espera-se um protótipo funcional com rastreabilidade auditável em escala completa e base técnica replicável para integração de matrizes curriculares em geração automática de itens, preservando julgamento especializado.

PROJETO 31**Orientador: Rafael Ferreira Leite de Mello****Título: Desenvolvimento de um sistema baseado em LLMs para correção de redações no ensino fundamental****Vagas: 1****Resumo:**

A inteligência artificial, por meio dos Grandes Modelos de Linguagem (LLMs), tem sido aplicada a tarefas educacionais, entre elas a correção automática de redações. No Ensino Fundamental, a correção de textos é uma atividade exigente para o professor, e o grande volume de produções dificulta devolutivas frequentes e individualizadas aos estudantes. O objetivo deste trabalho é projetar e desenvolver um sistema baseado em LLMs para apoiar a correção de redações nos anos finais do Ensino Fundamental, com rubricas alinhadas à BNCC e mecanismos de rastreabilidade, respondendo ao problema de como construir essa plataforma como apoio, e não substituto, do professor. Trata-se de pesquisa aplicada, descritiva e mista, com prototipação de uma API avaliada em três rodadas independentes sobre um corpus fixo de redações. Espera-se um protótipo funcional, registros de execução e relatório de avaliação técnica. Conclui-se que o sistema oferece base técnica auditável para a correção assistida, preservando a supervisão docente.

PROJETO 32**Orientador: Rafael Ferreira Leite de Mello****Título: Conformal Risk Control and Uncertainty Quantification for Reliable Large Language Models in Educational Assessment and Grounded Generation****Vagas: 1****Resumo:**

A crescente utilização de Large Language Models (LLMs) em sistemas educacionais tem ampliado suas aplicações em geração de respostas, explicações, feedbacks e avaliação automática. Entretanto, esses

modelos podem produzir saídas incorretas, pouco confiáveis ou não sustentadas pelas evidências disponíveis, tornando necessária a investigação de métodos capazes de quantificar e controlar sua incerteza. Este projeto visa investigar técnicas de quantificação de incerteza e controle estatístico de risco para LLMs em diferentes tarefas educacionais, buscando desenvolver mecanismos capazes de estimar a confiabilidade das saídas e determinar, de forma estatisticamente controlada, quando o sistema deve responder, realizar uma avaliação ou abster-se. Serão investigadas técnicas como Retrieval-Augmented Generation (RAG), Conformal Prediction, Conformal Risk Control, calibração, selective prediction e análise de incerteza semântica, incluindo a avaliação da fundamentação das respostas em nível de claim. Os métodos serão analisados em tarefas de geração e avaliação automática, como question answering, avaliação de textos e respostas abertas e avaliação de soluções matemáticas, considerando também cenários de mudança de distribuição e mecanismos adaptativos de controle de risco. A avaliação utilizará métricas de acurácia, calibração, cobertura, risco, taxa de abstenção e qualidade da fundamentação, visando contribuir para o desenvolvimento de LLMs educacionais mais confiáveis, calibrados e estatisticamente controlados.

PROJETO 33

Orientador: Rafael Ferreira Leite de Mello

Título: Feedback Formativo com LLMs e Gamificação personalizada

Vagas: 1

Resumo:

O feedback formativo é elemento central da aprendizagem, mas sua eficácia depende do engajamento ativo do estudante com o conteúdo recebido. Modelos de linguagem de grande escala (LLMs) têm emergido como alternativa para auxiliar na geração de feedback personalizado em contextos educacionais com turmas de diferentes dimensões. Contudo, a qualidade do feedback varia conforme o modelo utilizado, e a simples disponibilização do feedback gerado por IA mostra-se insuficiente para produzir efeitos de aprendizagem., uma vez que estudantes em ambientes digitais frequentemente ignoram ou leem superficialmente o feedback recebido. A gamificação, por sua vez, tem demonstrado potencial para aumentar o engajamento em ambientes educacionais, mas sua aplicação específica à entrega de feedback formativo, sobretudo de forma personalizada ao perfil do estudante, permanece pouco explorada. Neste contexto, este projeto investiga como LLMs podem ser utilizados para gerar feedback formativo de qualidade e em que medida a gamificação da entrega desse feedback, adaptada ao perfil dos estudantes, promove o engajamento com o conteúdo recebido.

PROJETO 34

Orientador: Ricardo André Cavalcante de Souza

Título: Extensão de Arquitetura Neuro-Simbólica para Grafos do Conhecimento, aplicada a domínios diversos

Vagas: 3

Resumo:

A Engenharia do Conhecimento baseada em Ontologias constitui uma abordagem para aquisição, organização, representação, integração e exploração formal do conhecimento de determinado domínio. Nessa perspectiva, ontologias descritas por meio de linguagens formais, como OML/OWL, permitem especificar conceitos, propriedades, relações, restrições e regras do domínio, estabelecendo uma camada

semântica comum para integrar dados provenientes de fontes diversas e heterogêneas. As instâncias de uma Ontologia podem ser materializadas na forma de Grafos de Conhecimento, representados no padrão RDF, os quais são armazenados em bases de dados semânticas. Dessa forma, dados originalmente dispersos podem ser transformados em conhecimento estruturado, interoperável e semanticamente relacionado, possibilitando identificar relações, inferências e correlações que subsidiem a geração de métricas, indicadores, análises, infográficos e outros instrumentos de apoio à tomada de decisão. A Engenharia do Conhecimento baseada em Ontologias pode ser apoiada por uma Arquitetura Neuro-Simbólica. O componente simbólico, representado pela Ontologia, RDF e mecanismos de raciocínio semântico, é responsável pela representação explícita e formal do conhecimento, definição das relações entre os conceitos e realização de inferências logicamente fundamentadas. Os Grafos de Conhecimento funcionam como elemento de integração entre os componentes da arquitetura, materializando entidades, propriedades e relações provenientes das diferentes fontes de dados de acordo com a semântica estabelecida pela Ontologia. O componente neural, representado por Large Language Models (LLMs), é responsável principalmente pela interpretação de conteúdos estruturados, semiestruturados e não estruturados e pela identificação e extração automatizada ou semiautomatizada de entidades, atributos e relações que possam ser utilizadas na população dos Grafos de Conhecimento. Estabelece-se, assim, uma relação complementar: enquanto os modelos neurais oferecem capacidade de interpretação, generalização e processamento de grandes volumes de informação heterogênea, os componentes simbólicos fornecem estrutura, semântica explícita, rastreabilidade e mecanismos formais de raciocínio sobre o conhecimento produzido. As oportunidades de pesquisa situam-se na investigação de como essa arquitetura neuro-simbólica pode ser estendida para:

- (1) Proporcionar maior contextualização, qualidade, consistência e confiabilidade ao processo de geração e exploração dos Grafos de Conhecimento. Para isso, propõe-se investigar a incorporação de Retrieval-Augmented Generation (RAG) e SHACL (Shapes Constraint Language) à arquitetura;
- (2) Comparação da eficiência e eficácia entre LLM e Small Language Models (SLM) como componentes neurais da arquitetura;
- (3) Aplicar e avaliar abordagem proposta em domínios diversos, desde avaliação de políticas públicas baseada em evidências, até apoio à tomada de decisão em atividades de sistemas agroalimentares como Avicultura e Pesca.

Esta proposta de pesquisa está fortemente alinhada com políticas de desenvolvimento econômico e social, tais como: Nova Indústria Brasil (NIB) que estabelece entre suas missões o desenvolvimento de cadeias agroindustriais sustentáveis e digitais; Estratégia Nacional de Ciência, Tecnologia e Inovação (ENCTI 2024–2034), em seu eixo relacionado à integração e fortalecimento do Sistema Nacional de CT&I; e Plano Brasileiro de Inteligência Artificial (PBIA 2024–2028), que estabelece ações estruturantes relacionadas à infraestrutura e ao desenvolvimento de IA, formação de capacidades, utilização de IA para melhoria dos serviços públicos e inovação nos sistemas produtivos.

PROJETO 35

Orientador: Rodrigo Lins Rodrigues

Título: Fatores associados à superação educacional em resolução de problemas computacionais entre estudantes socioeconomicamente vulneráveis

Vagas: 1

Resumo:

Este projeto propõe o desenvolvimento de modelos computacionais preditivos e explicáveis para a identificação da resiliência acadêmica de estudantes, com foco na nova base de dados do PISA 2025, cujo domínio inovador, Learning in the Digital World (LDW), avalia a capacidade de construir conhecimento e resolver problemas em ambientes digitais de aprendizagem, articulando resolução de problemas computacionais e autorregulação da aprendizagem. A entrada desse domínio no PISA amplia a própria noção de superação escolar: além da resiliência tradicional (estudantes do quartil inferior do índice ESCS que alcançam desempenho elevado nos domínios cognitivos), o projeto investigará a resiliência digital, isto é, estudantes socioeconomicamente vulneráveis e com acesso restrito a tecnologia que ainda assim demonstram alto desempenho em resolução de problemas computacionais e estratégias eficazes de autorregulação. Serão aplicados e comparados algoritmos supervisionados de aprendizado de máquina (Random Forest, XGBoost, SVM) e abordagens estatísticas de referência (regressão logística e modelos multiníveis com efeitos aleatórios de escola), respeitando o desenho amostral complexo do PISA por meio de valores plausíveis e pesos replicados de Fay/BRR. À modelagem preditiva será somada uma camada de explicabilidade com técnicas de XAI (SHAP e DALEX), a fim de identificar quais condições de escola, de família e de uso pedagógico de tecnologia mais contribuem para a superação em contextos de vulnerabilidade, produzindo resultados interpretáveis por formuladores de políticas públicas. O desenvolvimento se dará no ecossistema da linguagem R (tidymodels, EdSurvey, BIFIEsurvey, intsvy, WeMix, DALEX) com uso de sparklyr para processamento distribuído dos microdados. Bases complementares (TIMSS e SAEB) serão utilizadas para validação externa dos padrões encontrados e para ancorar as análises na realidade brasileira. Como contribuição, espera-se entregar um arcabouço metodológico e computacional reproduzível que permita identificar estudantes resilientes no contexto digital, dimensionar o peso da desigualdade de acesso e uso de tecnologia sobre essa superação e apoiar decisões de política educacional em um cenário de digitalização acelerada e uso crescente de inteligência artificial pelos estudantes.

Referências

- OECD (2026). PISA 2025 Assessment and Analytical Framework. Paris: OECD Publishing. <https://doi.org/10.1787/86c36975-en>
- OECD (2026). PISA 2025 Results (Volume I): Future-Ready Students. Paris: OECD Publishing. <https://doi.org/10.1787/73451bc5-en>
- Zhang, Y.; Cutumisu, M. (2024). Predicting the Mathematics Literacy of Resilient Students from High-performing Economies: A Machine Learning Approach. Studies in Educational Evaluation.
- Zheng, J.; Zhao, X.; Guo, X.; Chen, J.; Lu, S. (2025). Unfolding Key Factors of Resilience in ICT Cognitive-Motivational Engagement: Global Evidence from Machine Learning Techniques. International Journal of Educational Research.
- Álvarez-García, M.; Martínez-Córcoles, M.; Tordera, N.; Llorens, S. (2024). Uncovering Student Profiles: An Explainable Cluster-based Analysis Using PISA 2022 Data. Computers & Education.

PROJETO 36

Orientador: Rodrigo Lins Rodrigues

Título: Modelagem preditiva e explicável de perfis de uso e de concepções de inteligência artificial em dados educacionais heterogêneos

Vagas: 1

Resumo:

Estudantes brasileiros estão entre os que mais recorrem a ferramentas de inteligência artificial para estudar, o que configura adoção massiva e precoce em um sistema educacional que ainda não incorporou a temática ao currículo obrigatório. Familiaridade prática, contudo, não implica compreensão de mecanismos: a literatura mostra que estudantes descrevem sistemas como entidades que decidem por conta própria e confundem automação pré-programada com aprendizado de máquina (Lin & Dai, 2026), e que suas concepções permanecem ancoradas em programação e robótica mesmo após cursarem disciplinas de IA (Gao et al., 2024), ainda que a interação direta com sistemas de aprendizado de máquina permita raciocinar sobre a relação entre dados de treinamento e comportamento do modelo (Vartiainen et al., 2020). O projeto tem por objetivo ser desenvolvido em duas fases, na primeira, será utilizada os dados PISA 2025, que mediu pela primeira vez o uso de inteligência artificial e introduziu o domínio inovador Learning in the Digital World, onde será aplicado análise de classes latentes, algoritmos supervisionados comparados a modelos de regressão e técnicas de explicabilidade, respeitando o desenho amostral complexo por meio de valores plausíveis e pesos replicados de Fay/BRR. Na segunda fase, será desenvolvido um estudo de campo em escolas brasileiras, com coleta de dados multimodais por desenhos, atividade de ordenação de objetos por grau de inteligência, justificativas escritas e discussões em grupo transcritas, processados por um pipeline que combina análise de rede epistêmica, representações vetoriais contextuais e aprendizado supervisionado, validado contra codificação humana independente. A integração das fases consiste em prever a sofisticação conceitual a partir dos perfis de uso e das características do estudante. Esperam-se como contribuições um classificador validado para dados multimodais escolares em português, evidência empírica sobre a dissociação entre uso e compreensão sobre inteligência artificial e um pipeline aberto de modelagem preditiva e explicável sobre dados educacionais heterogêneos.

Referências

- OECD (2026). PISA 2025 Results (Volume I): Future-Ready Students. Paris: OECD Publishing. <https://doi.org/10.1787/73451bc5-en>
- OECD (2026). PISA 2025 Assessment and Analytical Framework. Paris: OECD Publishing. <https://doi.org/10.1787/86c36975-en>
- Lin, Z.; Dai, Y. (2026). Children's Conceptions of AI, Ethics and Intelligence in China: Evidence from Drawing and Ranking Activities. *British Journal of Educational Technology*, 57, 299-320.
- Bewersdorff, A.; Zhai, X.; Roberts, J.; Nerdel, C. (2023). Myths, Mis- and Preconceptions of Artificial Intelligence: A Review of the Literature. *Computers and Education: Artificial Intelligence*, 4, 100143. <https://doi.org/10.1016/j.caeai.2023.100143>
- Gao, H.; Zhang, Y.; Hwang, G.-J.; Zhao, S.; Wang, Y.; Wang, K. (2024). Delving into Primary Students' Conceptions of Artificial Intelligence Learning: A Drawing-Based Epistemic Network Analysis. *Education and Information Technologies*, 29, 25457-25486. <https://doi.org/10.1007/s10639-024-12847-0>
- Lin, Z.; Dai, Y. (2026). Children's Conceptions of AI, Ethics and Intelligence in China: Evidence from Drawing and Ranking Activities. *British Journal of Educational Technology*, 57, 299-320. <https://doi.org/10.1111/bjet.70016>
- Vartiainen, H.; Tedre, M.; Valtonen, T. (2020). Learning Machine Learning with Very Young Children: Who Is Teaching Whom? *International Journal of Child-Computer Interaction*, 25, 100182. <https://doi.org/10.1016/j.ijcci.2020.100182>

PROJETO 37

Orientadora: Taciana Pontual da Rocha Falcão

Título: Letramento Computacional no contexto de uma Educação Crítica em Computação

Vagas: 2

Resumo:

O Pensamento Computacional (PC) é uma habilidade de resolução de problemas baseada em fundamentos da computação como abstração, decomposição, identificação de padrões e algoritmos, que vem se consolidando como essencial para todos os cidadãos (BARR; STEPHENSON, 2011; CSTA/ISTE, 2011). Porém, para além dos aspectos cognitivos, existem abordagens que abordam o PC em um contexto de desenvolvimento consciente e crítico de tecnologias, democratizando não só o acesso mas também a produção de artefatos tecnológicos e os contextos sócio-econômicos afetados por ela. Essa perspectiva se alinha aos enfoques situado e crítico de PC propostos por Kafai et al. (2019): o PC situado é visto dentro do contexto daqueles que estão envolvidos em projetos de desenvolvimento de tecnologias, cujos interesses devem ser levados em consideração; e o PC crítico visa a justiça e equidade social, tendo a computação como ferramenta para evolução da humanidade e sua qualidade de vida. O PC crítico conecta-se à perspectiva da educação crítica em computação discutida por Ko et al. (2020). Segundo as pessoas autoras, a computação encontra-se no centro de muitas das crises globais como mudanças climáticas, guerras e pandemias. Em uma era em que a comunicação humana é primordialmente mediada por computadores e uma quantidade crescente de processos de produção é automatizada, a internet e as redes sociais amplificam notícias falsas, e algoritmos reforçam o racismo, sexismo, xenofobia, e outras formas de iniquidade e injustiça. Assim, ao colocar a computação como uma ferramenta para mudança social e resistência a forças opressoras (KAFAI et al., 2019), é preciso compreender que a computação não é, de forma alguma, neutra, mas influenciada por forças sociais, culturais, institucionais e políticas (FEENBERG, 2010, KO et al., 2020). A escolha de que tecnologia desenvolver envolve questões coletivas, morais e éticas que precisam ser discutidas, desenvolvendo um letramento crítico em computação e democratizando o desenvolvimento de tecnologias (KO et al., 2020). Este projeto visa investigar como processos participativos de desenvolvimento de artefatos digitais em contextos não hegemônicos podem ajudar a combater a exclusão digital por meio da democratização não somente do uso, mas também da produção de tecnologias.

Referências:

BARR, Valerie; STEPHENSON, Chris. 2011. Bringing computational thinking to K-12: what is involved and what is the role of the computer science education community? *Inroads*, v. 2, n. 1, p. 48-54. <https://doi.org/10.1145/1929887.1929905>

CSTA/ISTE. 2011. Computational Thinking: leadership toolkit. Disponível em: <https://c.ymcdn.com/sites/www.csteachers.org/resource/resmgr/471.11CTLeadershipToolkit-S.pdf>

FEENBERG, A. 2010. O que é a filosofia da tecnologia? *Racionalização Subversiva: Tecnologia, Poder e Democracia*, Observatório do Movimento pela Tecnologia Social na América Latina / CDS / UnB / Capes, Brasília.

KAFAI, Yasmin; PROCTOR, Chris; LUI, Debora. 2019. From Theory Bias to Theory Dialogue: Embracing Cognitive, Situated, and Critical Framings of Computational Thinking in K-12 CS Education. In: *Proceedings of the 2019 ACM Conference on International Computing Education Research*, p. 101-109. <https://doi.org/10.1145/3291279.3339400>

KO, A. J., OLESON, A., RYAN, N., REGISTER, Y., XIE, B., TARI, M., ... & LOKSA, D. (2020). It is time for more critical CS education. *Communications of the ACM*, 63(11), 31-33.

PROJETO 38

Orientadora: Taciana Pontual da Rocha Falcão

Título: Exploração de Novas Formas de Interação com Interfaces de Comunicação Alternativa

Vagas: 2

Resumo:

Um dos grandes desafios propostos para a comunidade brasileira de Interação Humano-Computador (IHC) em 2012 foi “Acessibilidade e Inclusão Digital” (FURTADO et al., 2015), centrado no acesso adequado de todas as pessoas aos sistemas digitais. Mais de uma década depois, apesar das diretrizes que buscam garantir a acessibilidade e a inclusão digital de PcD, muitos sistemas ainda são desenvolvidos sem considerar a diversidade das pessoas usuárias (PARTARAKIS; ZABULIS, 2024). Por outro lado, desde que os grandes desafios foram pensados em 2012, vivemos grandes avanços na ubiquidade dos dispositivos e na Inteligência Artificial (IA), cuja combinação traz um grande potencial para a inclusão não só digital, mas também social. Ainda hoje, as Pessoas com Deficiências (PcD) são excluídas das interações sociais, seja por falta de acessibilidade arquitetônica, seja por dificuldade de comunicação. A situação é ainda mais grave para pessoas não-verbais e com múltiplas deficiências, que não conseguem se comunicar oralmente e também podem possuir limitações motoras. Embora existam tecnologias que permitem que essas pessoas controlem uma máquina pelo piscar da pálpebra, eye-tracking ou ondas cerebrais, a efetiva inclusão delas na sociedade ainda é muito limitada, por questões de custo e portabilidade. Em aplicativos populares de Comunicação Alternativa e Aumentada baseados em seleção de imagens (Picture Exchange Communication System), tipicamente o conteúdo é criado por uma outra pessoa, o que limita o vocabulário (e consequentemente o poder de expressão) da PcD. Já os modelos usados pelos softwares de comunicação associados a sintetizadores de voz têm evoluído, por exemplo usando técnicas de processamento de linguagem natural (PLN) para predição de palavras. Entretanto, em ambos os casos persiste o problema da lacuna de reciprocidade (tempo de espera entre as falas dos interlocutores), que limita o grau de naturalidade em um diálogo. Uma comunicação que permita uma participação social efetiva precisa ser rápida e exigir pouco esforço, atributos de qualidade consolidados em IHC. Os avanços em IA Gerativa e Interativa (CHANG et al., 2023) têm um grande potencial para aumentar a autonomia e participação das PcD, viabilizando sua comunicação onde quer que estejam. Entretanto, historicamente, as inovações tecnológicas surgem com grande apelo no mundo hegemônico daqueles que possuem um desenvolvimento típico, sem considerar as PcD. Além de projetar dispositivos e interfaces que respondam a formas alternativas de interação, outra questão-chave é a produção de conteúdo, onde a IA Gerativa tem um papel fundamental. Os modelos de IA Gerativa são capazes de criar conteúdo, como textos, vídeos e imagens, a partir da identificação de padrões em grandes massas de dados, ou a partir de dados contextuais, como geolocalização e cena acústica. Assim, a IA Gerativa pode aumentar exponencialmente o espaço de escolha na comunicação da PcD, ampliando o vocabulário. Indo mais além, podemos pensar em modelos e interfaces que permitam que essas pessoas interajam diretamente com o sistema, criando seus próprios conteúdos através da IA, sem o auxílio de terceiros. Assim, é importante pensar em soluções inovadoras, combinando técnicas de IA a formas alternativas de interação, que tenham foco na inclusão social das PcD, viabilizando sua comunicação a partir de criação de conteúdo e assim permitindo sua participação efetiva e autônoma em contextos educacionais, profissionais e lúdicos.

Referências:

CHANG, Y., WANG, X., WANG, J., WU, Y., YANG, L., ZHU, K., ... & XIE, X. A survey on evaluation of large language models. ACM Transactions on Intelligent Systems and Technology, 2023.

FURTADO, E.S., CHAGAS, D., BITTENCOURT, I.I., FAÇANHA, A., Accessibility and Digital Inclusion. In: Baranauskas, de Souza and Pereira (orgs.). I GrandIHC-BR – Grand Research Challenges for Human-Computer Interaction in Brazil. Human-Computer Interaction Special Committee (CEIHC) of the Brazilian Computer Society (SBC). ISBN: 9788576692966. pp. 20-23. 2015.

PARTARAKIS, N.; ZABULIS, X. A Review of Immersive Technologies, Knowledge Representation, and AI for Human-Centered Digital Experiences. Electronics 2024, 13, 269. <https://doi.org/10.3390/electronics13020269>.

PROJETO 39

Orientador: Tiago Alessandro Espínola Ferreira

Título: Previsão e Análise de Séries Temporais Contínuas com o Uso de Redes Neurais Artificiais

Vagas: 1

Resumo:

Uma vasta gama de fenômenos do mundo real é um processo dinâmico, descrito por meio de uma série temporal. Uma série temporal é comumente representada por uma sequência de pontos ordenados sobre algum índice, geralmente o tempo. Contudo, em muitos problemas reais, o fenômeno que gera tal sequência de observações é descrito como um processo contínuo no tempo, sendo suas observações pontuais uma amostra do fenômeno. As equações diferenciais são uma possível modelagem matemática para a descrição de muitos destes fenômenos contínuos, sendo a solução das mesmas uma função que descreve a série temporal observada. Desta forma, séries temporais estocásticas contínuas podem ser modeladas por equações diferenciais estocásticas. Contudo, para muitos sistemas físicos/mundo real, de interesse prático, estas equações diferenciais estocásticas são analiticamente intratáveis. Nestes casos, existe um grande esforço de pesquisa no desenvolvimento de técnicas e procedimentos computacionais para resolver numericamente tais equações. A ideia central deste projeto de pesquisa é o estudo e o desenvolvimento de ferramentas computacionais baseadas em redes neurais artificiais (RNA), aqui em particular redes do tipo GAN (Generative Adversal Network) para a solução de equações diferenciais estocásticas dependentes do tempo, abordando o problema de previsão e análise de séries temporais no domínio temporal contínuo. Estará sendo aplicado o princípio básico de um problema de otimização. Definindo uma equação diferencial genericamente como $D(u) - F = 0$, onde D é a equação diferencial de interesse, u é uma possível solução de D e F é uma função conhecida. Seja $\hat{u}$ a saída de uma RNA. Se $\hat{u}$ é uma solução tentativa de D , então o resíduo da solução tentativa pode ser dado por $R(\hat{u}) = D(\hat{u}) - F$. A ideia básica é utilizar $R(\hat{u})$ ao quadrado como a função de erro (loss function) no processo de treinamento da RNA, onde o problema de resolução da equação diferencial é reduzido a um problema de minimização. Como caso de estudo, serão analisadas equações diferenciais estocásticas dependentes do tempo cujas soluções possam ser vistas como séries temporais contínuas de sistemas reais (mundo real).

PROJETO 40

Orientador: Tiago Alessandro Espínola Ferreira

Título: Desenvolvimento de Algoritmos Quânticos de Busca Baseados em Caminhadas Quânticas

Vagas: 1

Resumo:

Por meio da promessa de aumento da capacidade de processamento de informação, a computação quântica vem se destacando em alguns problemas específicos, onde a computação clássica (ou

convencional) apresenta alto custo computacional. Um destes tópicos é o “problema de busca em bases desordenadas”, onde a técnica de caminhadas quânticas vem encabeçando o desenvolvimento de várias abordagens para o estabelecimento de algoritmos quânticos de busca que são mais eficientes do que seus análogos clássicos. Um dos algoritmos baseados em caminhadas quânticas que apresenta grande interesse na literatura é a caminhada quântica baseada em estrutura com auto-laços (lackadaisical quantum walk). Neste sentido, este projeto de pesquisa visa o estudo e a análise de tais algoritmos, em particular com interesse em processos práticos de busca, como, por exemplo, o treinamento de redes neurais artificiais. Vários parâmetros necessitam de caracterização, como: a dependência das probabilidades de sucesso de busca em função da quantidade de autolaços, o tipo de estrutura balizadora para a definição do espaço de busca (grafos, grades, matrizes, listas, etc.), a distribuição e acoplamento dos arranjos espaciais das soluções buscadas, dentre outros. Desta forma, o objetivo deste projeto é o estudo e a análise dos parâmetros e estruturas para a evolução temporal do algoritmo de caminhadas quânticas com autolaços, com a intenção de desenvolver uma biblioteca em Python como uma ferramenta computacional para a elaboração de simulações de processos quânticos em computadores clássicos aplicadas a problemas de busca.

PROJETO 41

Orientador: Valmir Macario Filho

Título: Sistema Inteligente de Apoio ao Professor: Visão Computacional para Avaliação e Criação de Conteúdo de Matemática e Redação

Vagas: 1

Resumo:

Corrigir matemática é diferente de conferir resposta. O aluno erra o sinal ao passar um termo para o outro lado da equação e segue somando corretamente até o fim: a resposta está errada, mas o erro ocorreu três linhas antes. Ler isso exige acompanhar a estrutura da expressão, símbolo a símbolo. Na redação, o problema é o mesmo em outra escala: no ENEM, milhões de textos por edição, corretores diferentes, um único critério a aplicar. O projeto ataca o primeiro obstáculo de qualquer sistema desse tipo: o reconhecimento de manuscrito. São duas vagas, uma para expressões matemáticas e outra para redações, tendo como ponto de partida redes como SAN e CAN, que modelam a estrutura bidimensional das expressões, e TrOCR e HTR-Flor para texto corrido. Nenhuma delas foi pensada para a letra de um aluno brasileiro escrevendo numa folha de papel. A pesquisa está em adaptar e otimizar essas redes: alterações de arquitetura, como camadas adicionais de atenção e decodificadores sensíveis à sintaxe da expressão, e técnicas de treinamento, como aprendizado por transferência, data augmentation e geração de dados sintéticos. Ainda assim, o modelo erra, dada a variabilidade da caligrafia e a amplitude de símbolos que o aluno pode escrever. O pós-processamento é etapa crítica: validação sintática da saída, correção sensível ao contexto via modelos de linguagem e estimativa de confiança, para que o sistema sinalize quando não sabe em vez de arriscar um palpite. A avaliação usa ExpRate para expressões e CER/WER para texto. O resultado é um componente de reconhecimento robusto o bastante para ser acoplado a corretores automáticos, tutores inteligentes e outras ferramentas educacionais.

Referências:

Chevtchenko, S. F. et al. (2024). An Optimization Approach for Elementary School Handwritten Mathematical Expression Recognition. AIED 2024. Communications in Computer and Information Science, vol 2151. Springer, Cham.

Li, M. et al. (2023). TrOCR: Transformer-based optical character recognition with pre-trained models. AAAI, 37(2), 2579-2587.

De Sousa Neto, A. F., Bezerra, B. L. D., Toselli, A. H., & Lima, E. B. (2020). HTR-Flor: A deep learning system for offline handwritten text recognition. SIBGRAPI 2020, 216-223.

PROJETO 42

Orientador: Valmir Macario Filho

Título: Generalização de Modelos de Visão Computacional em Imagens de Ultrassom de Mama

Vagas: 1

Resumo:

Um modelo treinado e testado com imagens de um mesmo hospital costuma apresentar 97–99% de acurácia na classificação de lesões mamárias em ultrassom. O número raramente sobrevive à mudança de hospital. Muda o aparelho, muda o protocolo de aquisição, muda quem opera a sonda, muda a população atendida — e o desempenho que parecia resolvido despenca. Isso importa porque o câncer de mama é o mais incidente entre as mulheres no Brasil e o ultrassom, por ser barato, portátil e sem radiação, é a modalidade com maior alcance na rede pública, justamente onde a variação de equipamentos é maior. Parte do problema é que a rede aprendeu a coisa errada. Uma imagem de ultrassom carrega muito mais do que a lesão: o padrão de ruído característico do transdutor, o ganho escolhido pelo operador, as réguas e setas deixadas na tela, a moldura da interface do aparelho. Se essas marcas se correlacionam com o diagnóstico em uma base, o modelo as usa, e ninguém percebe enquanto o teste vier do mesmo hospital. O foco do projeto está nas pontas do pipeline. Na entrada, tratar a imagem antes que ela chegue à rede: recortar o campo de visão útil, remover anotações e marcadores, uniformizar o ganho, a escala e a resolução, e lidar com o speckle. Na saída, tratar a confiança do modelo: calibrar as probabilidades e dar a ele a opção de recusar um caso que não se parece com nada que viu, em vez de errar com convicção — o que, em diagnóstico, é o pior tipo de erro. Quando essas etapas indicarem que o gargalo está na própria representação aprendida, o trabalho avança para a arquitetura e o treinamento. O que se busca é desempenho que sobreviva ao mundo real: modelos avaliados em hospitais e equipamentos ausentes do treinamento, e não apenas na base em que foram desenvolvidos.

Referências:

Yu, Y., & Acton, S. T. (2002). Speckle reducing anisotropic diffusion. IEEE Transactions on Image Processing, 11(11), 1260-1270.

Gómez-Flores, W., Gregorio-Calas, M. J., & Pereira, W. C. A. (2024). BUS-BRA: A breast ultrasound dataset for assessing computer-aided diagnosis systems. Medical Physics, 51(4), 3110-3123.

Wang, L. (2026). Breast Ultrasound AI Under Dataset Shift: A Patient-Leakage-Aware Benchmark. Diagnostics, 16(10), 1537.