

Projetos Finais de Graduação para o 2s2024 (MC030)

Professor Responsável: [Prof. Dr. Bruno Barbieri de Pontes Cafeo](#)

Título: Explorando Técnicas de Mineração de Processos em Code Samples para Aprimorar o Desenvolvimento de Software

Descrição do Projeto: Este projeto propõe a aplicação de técnicas de mineração de processos em code samples. Code Samples são projetos completos de software disponibilizados por grandes organização (p.ex., Google, Amazon e Microsoft) para auxiliar desenvolvedores externos no aprendizado e atualização de conceitos implementados em seus produtos (p.ex., Android, AWS e Azure). Inspirados pelo sucesso do process mining, que utiliza a ciência de dados para descobrir, validar e melhorar workflows organizacionais, buscamos adaptar essa abordagem para o contexto de code samples. Os code samples servirão como unidades de análise, permitindo-nos extrair padrões, melhores práticas e áreas de otimização no processo de criação e evolução desses artefatos.

Trabalho: Individual

Professor Responsável: [Prof. Dr. Bruno Barbieri de Pontes Cafeo](#)

Título: Estudo Exploratório e Longitudinal das Práticas de Testes de Software em Code Samples

Descrição do Projeto: Este projeto propõe um estudo detalhado para explorar as práticas de testes de software em code samples, visando uma compreensão abrangente de como essas práticas evoluem ao longo do tempo. Code Samples são projetos completos de software disponibilizados por grandes organização (p.ex., Google, Amazon e Microsoft) para auxiliar desenvolvedores externos no aprendizado e atualização de conceitos implementados em seus produtos (p.ex., Android, AWS e Azure). Esses code samples serão analisados para identificar estratégias de teste, preferências de desenvolvedores e mudanças nas abordagens, proporcionando insights valiosos para aprimorar a eficácia dos testes nesse contexto que está em amplo crescimento.

Trabalho: Individual

Professor Responsável: [Prof. Dr. Bruno Barbieri de Pontes Cafeo](#)

Título: Explorando Technical Debt em Code Samples: Uma Análise Detalhada das Implicações para a Manutenibilidade do Código

Descrição do Projeto: Este estudo visa aprofundar a compreensão de technical debt (débito técnico) em code samples. Code Samples são projetos completos de software disponibilizados por grandes organização (p.ex., Google, Amazon e Microsoft) para auxiliar desenvolvedores externos no aprendizado e atualização de conceitos implementados em seus produtos (p.ex., Android, AWS e Azure). O technical debt refere-se a compromissos tomados durante o desenvolvimento que podem resultar em

custos adicionais no futuro. Ao analisar sua presença em code samples, buscamos identificar padrões, tipos comuns de technical debt e avaliar suas implicações na manutenibilidade do código desse tipo de artefato. Mais especificamente, este estudo colaborativo tem como objetivo realizar uma análise conjunta e inicial por meio de code smells e refatorações em code samples, com dois alunos se concentrando em áreas específicas. Enquanto um aluno se dedica à identificação e categorização de code smells, o outro foca na análise das refatorações existentes nos code samples. A proposta é, posteriormente, examinar a possível relação entre os code smells identificados e as refatorações realizadas, buscando entender como as práticas de refatoração podem influenciar na mitigação ou agravamento dos problemas identificados. Este enfoque integrado visa proporcionar insights mais abrangentes para o aprimoramento da qualidade dos code samples.

Trabalho: Em Equipe (discriminar na descrição a divisão dos trabalhos)

Professor Responsável: [Prof. Dr. Bruno Barbieri de Pontes Cafeo](#)

Título: Entendendo Práticas de Desenvolvimento de Code Samples: Uma Investigação por Meio de Entrevistas com Desenvolvedores

Descrição do Projeto: Este projeto visa explorar e compreender as práticas adotadas por desenvolvedores na criação e evolução de code samples, trechos de código-fonte que representam unidades significativas de implementação. Através de entrevistas com desenvolvedores atuantes em diferentes organizações, buscamos obter insights sobre as estratégias, desafios, e decisões tomadas durante o processo de criação, manutenção e evolução de code samples. O estudo proporcionará uma visão detalhada das práticas de desenvolvimento e permitirá identificar padrões e tendências no cenário de code samples.

Para este projeto é imprescindível um nível avançado de inglês para a condução das entrevistas, bem como posterior extração dos dados.

Trabalho: Individual

Professor Responsável: [Prof. Dr. Bruno Barbieri de Pontes Cafeo](#)

Título: Desenvolvimento de Métricas de Evolução em Sistemas Configuráveis Pre-processados utilizando o Modelo GQM

Descrição do Projeto: Este projeto propõe a aplicação do modelo Goal-Question-Metric (GQM) para a definição e implementação de métricas de evolução em sistemas configuráveis pre-processados. Sistemas configuráveis pre-processados são aqueles em que características específicas podem ser ativadas ou desativadas antes da compilação (p.ex., kernel do Linux, Mongo DB, entre outros). O GQM será utilizado como estrutura para estabelecer metas, identificar questões relevantes e, por fim, definir métricas objetivas que permitam avaliar e monitorar a evolução desses sistemas ao longo do tempo.

Trabalho: Individual

Professor Responsável: [Prof. Dr. Bruno Barbieri de Pontes Cafeo](#)

Título: Revisão Sistemática de Técnicas de Mineração de Padrões de Uso de APIs na Literatura

Descrição do Projeto: Este projeto propõe uma revisão sistemática para analisar as técnicas de mineração de padrões de uso de APIs presentes na literatura. O objetivo é compreender as abordagens utilizadas para identificar, extrair e analisar padrões de uso de APIs em diversos contextos. A pesquisa será conduzida em dupla, com ambas as partes colaborando em todas as etapas do processo, desde a revisão inicial da literatura até a análise final dos resultados. No final cada aluno realizará a análise dos dados de um conjunto de questões de pesquisa.

Trabalho: Em Equipe (discriminar na descrição a divisão dos trabalhos)

Professor Responsável: [Prof. Dr. Bruno Barbieri de Pontes Cafeo](#)

Título: Mineração de Padrões de Uso de APIs na Biblioteca Standard C

Descrição do Projeto: Este projeto visa aplicar técnicas consolidadas de mineração de dados para extrair padrões de uso de APIs na biblioteca Standard C da linguagem de programação C. Padrões de uso de APIs referem-se a práticas recorrentes adotadas pelos desenvolvedores ao utilizar interfaces de programação de aplicativos (APIs). Essas práticas incluem sequências comuns de chamadas de funções, estratégias de manipulação de dados, e abordagens para lidar com situações específicas. O objetivo é identificar comportamentos frequentes e tendências no uso das APIs padrão, oferecendo insights valiosos sobre práticas de desenvolvimento, desafios comuns e possíveis melhorias em documentação da biblioteca.

Trabalho: Individual

Professor Responsável: [Prof. Dr. Bruno Barbieri de Pontes Cafeo](#)

Título: Revisão de Literatura Multivocal sobre Feature Flags

Descrição do Projeto: Este projeto propõe iniciar uma revisão de literatura multivocal sobre feature flags, explorando diversas perspectivas e práticas relacionadas a essa abordagem no desenvolvimento de software. Feature flags, ou feature toggles, são mecanismos que permitem ativar ou desativar funcionalidades em tempo de execução. Esta revisão visa compreender a variedade de conhecimento existente na literatura formal e em sites, blogs, foruns, etc. sobre feature flags, abordando aspectos técnicos, práticos e conceituais. O projeto servirá como um ponto de partida para uma análise abrangente que capture as nuances e evoluções dessa prática no cenário de desenvolvimento de software.

Trabalho: Individual

Professor Responsável: [Prof. Dr. Carlos Alberto Astudillo Trujillo](#)

Título: Algoritmos RAG em LLMs Aplicados em Documentos sobre Redes Móveis 5G/6G

Descrição do Projeto: Este projeto propõe a implementação e avaliação de desempenho de algoritmos

RAG aplicados a documentos de redes móveis 5G/6G. Através da combinação de recuperação eficiente de informações e geração de texto, o objetivo é criar um sistema que possa fornecer respostas precisas e detalhadas às consultas baseadas nos padrões 3GPP, que definem as tecnologias de comunicações móveis 5G/6G. Isso é crucial para suportar o desenvolvimento e a implementação contínua das tecnologias de comunicações móveis, garantindo que as especificações sejam corretamente interpretadas e aplicadas.

Os principais desafios incluem a indexação eficiente dos documentos do 3GPP, a integração entre os modelos de recuperação e geração, e a avaliação rigorosa do desempenho do sistema.

Trabalho: Individual

Professor Responsável: [Prof. Dr. Carlos Alberto Astudillo Trujillo](#)

Título: Análise de Desempenho de Esquemas de Acesso Aleatório em Redes 5G/6G Integradas com Satélites de Órbita Terrestre Baixa (LEO) para Acesso Massivo de Dispositivos IoT

Descrição do Projeto: O objetivo deste projeto é estudar mecanismos de acesso aleatório em redes 5G/6G integradas com satélites de órbita terrestre baixa (LEO) para a Internet das coisas. As atividades deste projeto incluem a implementação e teste de um módulo de simulação para o procedimento de acesso aleatório neste tipo de redes. Será realizada também a avaliação de desempenho de mecanismos selecionados de acesso aleatório utilizando o simulador desenvolvido no projeto. Serão analisadas as implicações da integração com satélites LEO, considerando os desafios específicos dessa configuração, como a mobilidade do satélite e as características das trajetórias orbitais

Trabalho: Individual

Professor Responsável: [Prof. Dr. Carlos Alberto Astudillo Trujillo](#)

Título: Simulador de Constelações de Satélites de Órbita Terrestre Baixa (LEO) para Avaliação de Desempenho de Tecnologias de Comunicação Satelitais

Descrição do Projeto: O projeto de TCC propõe o desenvolvimento de um simulador para constelações de satélites em órbita terrestre baixa (LEO). O objetivo é fornecer uma ferramenta para poder ser incluída em simuladores de rede para avaliar o desempenho de tecnologias de comunicação satélite. O simulador permitirá a análise de diferentes configurações de constelações, bem como a otimização de parâmetros para garantir a eficiência e confiabilidade das redes de comunicação via satélite. A

ferramenta permitirá incluir as implicações da integração de redes terrestres com satélites LEO, assim como considerar os desafios específicos, como mobilidade dos satélites e as características das trajetórias orbitais.

Trabalho: Individual

Professor Responsável: [Prof. Dr. Carlos Alberto Astudillo Trujillo](#)

Título: Otimização e Avaliação de um Sistema de Aquisição de Dados para Experimentos de Física de Partículas

Descrição do Projeto: Este projeto visa otimizar e avaliar um sistema de aquisição de dados (DAQ, em Inglês) usando o chip SAMPa para detectores GEM (Gas Electron Multiplier) do experimento Alice (CERN/LHC). O foco é melhorar a integração do chip SAMPa, desenvolvido na USP, otimizado para processamento de sinal de baixa carga e alta resolução, com detectores GEM de forma a melhorar a precisão e eficiência na detecção de partículas.

Trabalho: Individual

Professor Responsável: [Prof. Dr. Carlos Alberto Astudillo Trujillo](#)

Título: Modelo de Consumo de Energia baseado em Aprendizado de Máquina para a Interface de Rede de Dispositivos 5G

Descrição do Projeto: Este projeto tem como objetivo desenvolver um modelo de consumo de energia para a interface de rede 5G/6G, especialmente direcionado a dispositivos para a Internet das Coisas (IoT). Utilizando técnicas de aprendizado de máquina, o modelo será capaz de estimar o consumo energético com base em parâmetros como quantidade de recursos utilizados, potência de transmissão, potência de recepção, largura de banda e características físicas específicas dos dispositivos, tais como número de antenas. A aplicação desse modelo permitirá otimizar o uso da interface de rede 5G/6G, prolongar a vida útil das baterias dos dispositivos IoT e contribuir para a eficiência energética em ambientes conectados.

Trabalho: Individual

Professor Responsável: [Prof. Dr. Carlos Alberto Astudillo Trujillo](#)

Título: Algoritmos de Decodificação Eficientes para Esquemas de Acesso Aleatório baseado em Código Expandido em Redes Móveis 5G/6G

Descrição do Projeto: O objetivo deste projeto é implementar e avaliar o desempenho de algoritmos de decodificação eficientes para esquemas de acesso aleatório baseado em acesso aleatório de código expandido (CeRA). No CeRA, cada dispositivo IoT executa acesso aleatório usando transmissões consecutivas de preâmbulos selecionados aleatoriamente em n slots de acesso aleatório (RA), formando um superframe. Em cada subframe de RA, um preâmbulo é selecionado aleatoriamente dentre um conjunto de preâmbulos

permitidos. Essa sequência de preâmbulos é interpretada como uma palavra de código de comprimento n , e recursos para transmissão de dados são alocados com base nas palavras-código inferidas pela estação rádio base no superframe.

CeRA é um dos esquemas mais inovadores e promissores para o suporte de dispositivos IoT em redes móveis 5G/6G. No entanto, ele também é desafiador porque requer algoritmo eficiente de decodificação. Esses algoritmos são empregados para inferir as palavras-código transmitidas na estação rádio base, baseado apenas nos preâmbulos detectados em cada um dos n slots de acesso aleatório que compõem um superslot. Os algoritmos de decodificação existentes são exponenciais em n . Assim, o problema de decodificação eficiente em CeRA precisa ser abordado para explorar plenamente o seu verdadeiro potencial em redes 5G/6G.

Trabalho: Individual

Professor Responsável: [Prof. Dr. Carlos Alberto Astudillo Trujillo](#)

Título: Gêmeos Digitais para Previsão de Falhas de Dispositivos IoT Industriais

Descrição do Projeto: Este projeto visa desenvolver um sistema de previsão de falhas para dispositivos IoT (Internet das Coisas) industriais utilizando a arquitetura de gêmeos digitais. O sistema permitirá correlacionar dados em tempo real dos dispositivos com informações de seus gêmeos digitais, permitindo antecipar possíveis problemas, como mau funcionamento, desgaste ou falhas iminentes. Isto ajudará na manutenção proativa, redução de custos e aumento da confiabilidade dos dispositivos IoT em aplicações de nível industriais.

Trabalho: Individual

Professor Responsável: [Prof. Dr. Carlos Alberto Astudillo Trujillo](#)

Título: Previsão de Demandas de Rede e Computação para Redes 5G/6G Integradas com Satélites de Órbita Baixa da Terra

Descrição do Projeto: O objetivo deste projeto é desenvolver um modelo baseado em aprendizado de máquina para previsão das demandas de rede e computação em aplicações de computação de borda móvel sobre redes 5G/6G integradas com satélites de baixa órbita terrestre (LEO).

Esse mecanismo será utilizado para realizar alocação eficiente dos recursos do sistema.

Trabalho: Individual

Professor Responsável: [Prof. Dr. Edmundo Roberto Mauro Madeira](#)

Título: Aprendizado de Máquina aplicado a alocação de recursos de rede

Descrição do Projeto: Infraestruturas de rede necessitam gerenciar de forma eficiente os seus recursos - capacidade de processamento, armazenamento, transmissão de dados, dentre outros - para atender as necessidades das aplicações e serviços

acessados por seus usuários. Devido a variações na demanda desses recursos - alteração no número de usuários, mobilidade dos usuários, ou alterações intrínsecas da própria aplicação, modificações nas configurações das redes se fazem necessárias. Este projeto tem como objetivo estudar a aplicação de soluções baseadas em aprendizado de máquina para identificar as alterações necessárias, ao longo do tempo, para realizar de forma eficiente a (re)distribuição dos recursos da rede dado suas variações na demanda.

Trabalho: Individual

Professora Responsável: [Pro^a. Dr^a. Eliane Martins](#)

Título: Estratégias de testes de imparcialidade para sistemas de recomendação

Descrição do Projeto: Sistemas de recomendação são utilizados para sugerir itens relevantes para os usuários, ao mesmo tempo em que ajudam aos provedores de conteúdo a auxiliar seu público alvo. São úteis para auxiliar usuários na busca quando há um grande volume de itens para analisar. São muito utilizados em comércio eletrônico, streaming, e-science, entre outros. Uma dificuldade com estes sistemas é que eles podem fazer recomendações tendenciosas, que podem resultar em discriminação de grupos de usuários ou itens. Testes para determinar a imparcialidade ou justiça (fairness) desse tipo de sistema ainda apresenta muitos desafios. Este projeto tem por objetivo estudar estratégias propostas para testes de imparcialidade de sistemas de recomendação. Este estudo visa antes de mais nada conhecer as diferentes definições de imparcialidade, bem como as estratégias e métricas utilizadas para fins de testes. Eventualmente, utilizar uma das estratégias estudadas nos testes de um estudo de caso.

Trabalho: Individual

Professor Responsável: [Prof. Dr. Hervé Cédric Yviquel](#)

Título: Implementação e Otimização de Modelos de IA com a Linguagem de Programação Mojo

Descrição do Projeto: Este projeto visa explorar a linguagem de programação Mojo, desenvolvida pela startup Modular, para implementar e otimizar modelos de inteligência artificial (IA), aproveitando a integração com a sintaxe e as bibliotecas do Python e a eficiência de sistemas de baixo nível. A escolha de Mojo deve-se à sua capacidade de executar algoritmos complexos com eficiência significativamente superior à do Python padrão, aproveitando características de hardware especializado como TensorCores e extensões AMX.

Trabalho: Individual ou Em Equipe (discriminar na descrição a divisão dos trabalhos)

Professor Responsável: [Prof. Dr. Hélio Pedrini](#)

Título: Reconhecimento de Ações em Vídeo.

Descrição do Projeto: Identificação de ações humanas em sequências de vídeos por meio de técnicas de processamento de imagens, análise de vídeos, visão computacional e aprendizado de máquina.

Trabalho: Individual

Professor Responsável: [Prof. Dr. Hélio Pedrini](#)

Título: Identificação de Eventos Anômalos em Vídeos de Vigilância.

Descrição do Projeto: Identificação de comportamento anômalo em vídeos de vigilância baseada em técnicas de visão computacional, análise de imagens e aprendizado de máquina.

Trabalho: Individual

Professor Responsável: [Prof. Dr. Hélio Pedrini](#)

Título: Análise de Imagens e Vídeos.

Descrição do Projeto: Proposição e aplicação de técnicas para processamento e análise de imagens e vídeos em diferentes domínios de conhecimento (sensoriamento remoto, medicina, biologia, biometria, microscopia, vigilância e segurança).

Trabalho: Individual

Professor Responsável: [Prof. Dr. Luiz Fernando Bittencourt](#)

Título: Aprendizado de máquina federado no simulador MobFogSim

Descrição do Projeto: O simulador MobFogSim simula a mobilidade de dispositivos de usuário que podem executar aplicações em dispositivos de processamento presentes na borda da rede, ou computação em névoa, complementando as funcionalidades da computação em nuvem. Este projeto tem como objetivo implementar novas funcionalidades no simulador, mais especificamente a implementação de aprendizado de máquina federado nos dispositivos do simulador.

Trabalho: Individual ou Em Equipe (discriminar na descrição a divisão dos trabalhos)

Professor Responsável: [Prof. Dr. Luiz Fernando Bittencourt](#)

Título: Elasticidade e auto-distribuição

Descrição do Projeto: Sistemas que se auto-distribuem (do inglês self-distributing systems) são sistemas capazes de replicar, em tempo de execução, componentes que

compõem sua própria estrutura, lidando com estado atrelado a esses componentes de forma transparente. Esses sistemas têm por objetivo explorar ambientes contemporâneos como computação em nuvem, que apesar de terem software que dão apoio a elasticidade, não lidam bem com a replicação de serviços com estado. Este projeto tem por objetivo expandir o conceito de auto-distribuição de componentes para explorar placement de serviços com estado da nuvem para a edge e vice-versa de forma transparente, podendo assim aproveitar as vantagens de ambos ambientes operacionais.

Trabalho: Individual ou Em Equipe (discriminar na descrição a divisão dos trabalhos)

Professor Responsável: [Prof. Dr. Luiz Fernando Bittencourt](#)

Título: Gerência de Sistemas Auto-adaptativos

Descrição do Projeto: A crescente complexidade na criação e gestão de sistemas distribuídos está centrada, principalmente, na volatilidade dos ambientes operacionais modernos. O ambiente volátil é caracterizado por constantes mudanças, muitas vezes inesperadas, que o sistema precisa lidar em tempo de execução para se manter funcional e atender seus requisitos não-funcionais. Para lidar com essas constantes mudanças, sistemas auto-adaptativos, capazes de se auto-adaptarem diante de mudanças, estão ganhando cada vez mais destaque. Este projeto tem como objetivo a exploração de algoritmos de aprendizado de máquina por reforço para aprender, em tempo de execução, a como adaptar e re-adaptar sistemas distribuídos em ambientes voláteis sem interferência humana, de forma a preservar sua funcionalidade ou otimizar aspectos não-funcionais do sistema.

Trabalho: Individual ou Em Equipe (discriminar na descrição a divisão dos trabalhos)

Professor Responsável: [Prof. Dr. Luiz Fernando Bittencourt](#)

Título: Aprendizado de Máquina Distribuído

Descrição do Projeto: Conjuntos de datasets são encontrados de diferentes variedades e gerados através de várias operações de upload ou offloading providas de dispositivos IoT ou de sistemas autônomos em aplicações Big Data streaming. O cálculo de algoritmos de Deep Learning (DL) computa um gradiente, chamado de Stochastic Gradient Descent (SGD). Este cálculo apresenta problemas diversos quando explorado de forma distribuída, dessa forma a aplicação de novas ferramentas para avaliar tais dados de forma distribuída é um grande desafio. Este projeto visa avaliar a performance de diferentes aplicações e algoritmos de DL, considerando acurácia, precisão, tempo de execução, número de épocas e quantidade de trocas de mensagens, em um ambiente de Internet das Coisas.

Trabalho: Individual ou Em Equipe (discriminar na descrição a divisão dos trabalhos)

Professor Responsável: [Prof. Dr. Luiz Fernando Bittencourt](#)

Título: Gerência de recursos em sistemas distribuídos

Descrição do Projeto: A gerência de recursos envolve o processo de seleção dos recursos computacionais para execução de aplicações de diversos tipos. A otimização da alocação de recursos, como no escalonamento de tarefas, depende do desempenho que tais tarefas obtêm do recurso computacional em questão, seja este de processamento, armazenamento ou de rede. Este projeto tem como objetivo identificar uma ou mais aplicações e um ambiente de processamento distribuído para realização de uma análise de questões que concernem a gerência de recursos, tais como formas de implementação da aplicação, análise de desempenho e algoritmos de alocação de recursos.

Trabalho: Individual ou Em Equipe (discriminar na descrição a divisão dos trabalhos)

Professor Responsável: [Prof. Dr. Luiz Fernando Bittencourt](#)

Título: Monitoramento de colmeias com Internet das Coisas

Descrição do Projeto: Este trabalho é parte de um projeto de extensão coordenado pelo Prof. Roberto Greco do IG. O objetivo é projetar e implementar um sistema IoT que usa arduínos e/ou raspberries e sensores para monitoramento de colmeias de abelhas nativas (e.g. temperatura, umidade, som).

Trabalho: Individual ou Em Equipe (discriminar na descrição a divisão dos trabalhos)

Professor Responsável: [Prof. Dr. Luiz Fernando Bittencourt](#)

Título: Sistemas distribuídos para realidade virtual e aumentada

Descrição do Projeto: O(s) estudante(s) se aprofundarão em problemas relacionados à execução de aplicações e serviços distribuídos de realidade virtual e aumentada. Problemas de transmissão e qualidade de vídeo e processamento em tempo real para dispositivos de borda e vestíveis serão estudados para proposição de soluções viáveis e com desempenho adequado.

Trabalho: Individual ou Em Equipe (discriminar na descrição a divisão dos trabalhos)

Professor Responsável: [Prof. Dr. Luiz Fernando Bittencourt](#)

Título: Prevenção de Colisão e Controle de Rotas de Drones em Cenários com Comunicação 5G/6G

Descrição do Projeto: Cenários onde múltiplos drones co-existem demandam sistemas de prevenção de colisão e estabelecimento de rotas eficientes. O controle dos drones pode ser feito de forma cooperativa, com troca de informações entre eles, ou não-cooperativa, onde desvios

e rotas são estabelecidos autonomamente por cada drone. Este tópico visa estudar e explorar estes cenários, através de simulações, onde a troca de dados pode ou não ocorrer através de pontos de acesso de rede celular 5G/6G.

Trabalho: Individual ou Em Equipe (discriminar na descrição a divisão dos trabalhos)

Professor Responsável: [Prof. Dr. Ruben Interian Kovaliova](#)

Título: Análise de redes de colaboração científica

Descrição do Projeto: Grupos de pesquisa são um componente chave dentro de Universidades e Institutos, e a sua formação e características possuem um alto impacto na qualidade da pesquisa e na formação de professores e pesquisadores.

Este projeto tem como objetivo a extração de conhecimento de grandes grafos de coautoria e de citações entre publicações científicas para identificar grupos de pesquisa no Brasil com diferentes graus de modularidade e integração com outros grupos de pesquisa nacionais e internacionais, utilizando a Unicamp especificamente como estudo de caso. O projeto visa contribuir para uma melhor tomada de decisão pela direção das diversas instituições acadêmicas, científicas e do governo, incluindo a própria Unicamp, e subsidiar políticas públicas eficientes de modo a estimular grupos que trabalham em áreas de pesquisa específicas a melhorar a sua inserção em redes internacionais de pesquisa, ou com outros pesquisadores no Brasil, encontrando ativamente formas de fomentar essa integração.

O trabalho envolverá a detecção de comunidades subjacentes dentro ou entre grupos de afiliação institucional dos autores, com diferentes graus de modularidade, usando algoritmos de clusterização hierárquica em grafos.

Trabalho: Individual

Professor Responsável: [Prof. Dr. Sandro Rigo](#)

Título: Simulador de Arquitetura para Aceleradores Matriciais

Descrição do Projeto: Operações Matriciais são muito importantes para cargas de trabalho de IA e HPC. Em particular, multiplicação de matrizes é a operação mais importante para aprendizado de máquinas. Por isso, ao longo dos últimos anos, o desenvolvimento de HW especializado para aceleração desse tipo de operação tem estado em evidência. Porém, existem várias decisões arquiteturais a serem tomadas que determinam o suporte que será dado por um acelerador em termos de formatos de matrizes oferecidos, tipos de dados, portabilidade de código, etc. O objetivo desse trabalho é criar um simulador onde essas decisões arquiteturais possam ser estudadas de forma precisa e visual. Com isso, o trabalho de tomada dessas decisões se torna mais claro, além de dar suporte ao ensino de forma que as consequências de cada escolha possam ser exploradas pelo aluno. A ideia é desenvolver o simulador em Python baseado no conjunto de instruções RISC-V.

Trabalho: Individual

Professor Responsável: [Prof. Dr. Sandro Rigo](#)

Título: Deploy de Modelos LLM em Sistemas Distribuídos

Descrição do Projeto: O objetivo desse projeto é estudar frameworks para escalar o treinamento de grandes modelos de ML em clusters. Em particular, estamos interessados em usar uma ferramenta como o t5X para escalar o treinamento de LLMs como Llama, qwen2 ou outros modelos que estejam disponíveis. Os experimentos iniciais serão executados em nosso cluster no LSC, mas podemos escalar para sistemas maiores no futuro. É possível termos um trabalho em equipe devido a variedade de modelos disponíveis.

Trabalho: Individual

Professor Responsável: [Prof. Dr. Zanoni Dias](#)

Título: Ordenação Semi-Completa por Rearranjo de Genomas

Descrição do Projeto: Problemas de Rearranjo de Genomas buscam estimar a distância evolutiva entre genomas de diferentes organismos. Estes problemas lidam com eventos de rearranjo, que são mutações capazes de alterar a sequência genética dos genomas. Quando assumimos que os genomas comparados não possuem genes repetidos, o problema corresponde a ordenação de uma permutação. O objetivo deste projeto consiste em desenvolver heurísticas que, dado um genoma de origem e um genoma alvo, sejam capazes de fornecer um genoma (obtido através do menor número de eventos de rearranjo de genomas) que esteja o mais próximo de genoma alvo possível. O problema ainda possui uma restrição sobre a proximidade máxima permitida para uma solução válida, que é estabelecida como entrada para o problema e considera as posições dos genes.

Trabalho: Individual

Professor Responsável: [Prof. Dr. Zanoni Dias](#)

Título: Problemas de Partição de Strings

Descrição do Projeto: Problemas de Partição de Strings estão intimamente ligados a problemas de Rearranjo de Genomas. Esses problemas consistem em segmentar duas strings, dadas como entrada do problema, de forma que as partes de uma das strings possam ser reorganizadas para obtermos a outra string. A forma como as strings podem ser segmentadas ou como as partes podem ser reorganizadas da origem formam diferentes variações do problema. Por exemplo, cada uma das partes pode ou não ser invertida ao realizarmos a reorganização ou podemos permitir que algumas partes sejam deletadas ou inseridas antes da reorganização. O objetivo deste projeto é o desenvolvimento de algoritmos exatos para problemas de Partição de Strings. Existem vários caminhos possíveis que podem ser usados para tais algoritmos, dentre eles temos algoritmos parametrizados, programação linear inteira e decomposição de grafos em ciclos.

Trabalho: Individual

Professor Responsável: [Prof. Dr. Zanoni Dias](#)

Título: Análise Comparativa de Algoritmos para Problemas de Rearranjo de Genomas

Descrição do Projeto: Os problemas de Rearranjo de Genomas permitem a comparação de genomas de diferentes indivíduos e têm múltiplas aplicações na área da Biologia Computacional. Como esses problemas já vêm sendo estudados há muitos anos, várias versões foram propostas levando em conta as características da representação dos genomas utilizada e as operações de rearranjo consideradas. Além disso, muitos algoritmos para o problema podem ser decompostos em diferentes etapas, cada uma podendo ser resolvida por múltiplos algoritmos diferentes. O objetivo deste projeto é realizar uma comparação experimental de diferentes abordagens conhecidas para comparar genomas com base em problemas de rearranjo. As comparações poderão ser realizadas em bases de dados geradas artificialmente e com dados de genomas reais.

Trabalho: Individual

Professor Responsável: [Prof. Dr. Zanoni Dias](#)

Título: Visualização e Interpretação de Modelos Profundos

Descrição do Projeto: Entender e significar o funcionamento de sistemas complexos, como algoritmos de aprendizado de máquina baseados em redes neurais artificiais profundas, são atividades empregadas nos estágios de avaliação, aplicação e sustentação de sistemas autônomos e necessárias para confirmação do funcionamento correto dos mesmos. O objetivo deste projeto é estudar as diferentes técnicas de visualização e explicabilidade para modelos profundos existentes na literatura, focando-se em um ou mais algoritmos modal-genéricos, como: técnicas baseadas em gradiente, Class Activation Mapping (CAM) e Layer-wise Relevance Propagation (LRP).

Trabalho: Individual

Professor Responsável: [Prof. Dr. Zanoni Dias](#)

Título: Segmentação Semântica Fracamente Supervisionada da Amazônia

Descrição do Projeto: Neste projeto, serão exploradas técnicas de segmentação semântica fracamente supervisionada —, como técnicas de transferência de conhecimento prévio, técnicas de regularização sobre rótulos fracos e Few-Shot Segmentation — a fim de criar anotações de segmentação semântica (ao nível de píxel) sobre o conjunto de dados Planet: Understanding the Amazon from Space, a partir de somente rótulos de classificação ruidosos (ao nível de imagem).

Trabalho: Individual

Professor Responsável: [Prof. Dr. Zanoni Dias](#)

Título: Desenvolvimento de ferramenta web para classificação de proteínas

Descrição do Projeto: A determinação das funções desempenhadas por cada proteína demanda um grande esforço técnico e financeiro. Com o intuito de auxiliar neste desafio, diversos métodos computacionais vêm sendo aplicados para a tarefa em questão. Diante deste cenário, surge a necessidade de desenvolver ferramentas online a partir das abordagens computacionais que possam auxiliar os pesquisadores no mapeamento das funções das proteínas, servindo como ponto de partida para as investigações laboratoriais. O objetivo deste projeto é o desenvolvimento de uma ferramenta web destinada à classificação de proteínas, a partir de uma abordagem de aprendizado de máquina já estabelecida.

Trabalho: Individual

Professor Responsável: [Prof. Dr. Zanoni Dias](#)

Título: Análise de redes de interações de proteínas

Descrição do Projeto: Determinar as funções exercidas por cada proteína requer um grande esforço técnico e financeiro. As redes de interações, que mapeiam as interações funcionais entre as proteínas, são ferramentas valiosas para resolver este problema. Através da análise dessas redes, é possível inferir a função de uma proteína desconhecida com base nas funções das proteínas com as quais ela interage. O objetivo deste projeto é empregar técnicas de análise de redes de interações e integrar essa abordagem com uma metodologia já existente para a análise de sequências de aminoácidos.

Trabalho: Individual