

Professor Responsável	Título	Descrição dos Projetos Finais de Graduação para o 1s2021	Individual/Equipe
Eliane Martins	Uso de fuzzers nos testes de robustez e segurança de bibliotecas de manipulação de imagens	Fuzz testing (ou fuzzing) é uma técnica de teste de software que insere dados aleatórios (válidos ou inválidos) no sistema de software para descobrir defeitos no tratamento de entradas, bem como brechas de segurança. Os dados são inseridos usando técnicas de teste automatizadas durante a execução, em que o sistema é monitorado para determinar se ocorrerem exceções, queda (crash) ou bloqueio (hang) do sistema. Neste projeto serão estudadas técnicas de fuzzing para dados de entrada estruturados de entrada, como os imagens. O objetivo é avaliar dois tipos de técnicas de fuzzing, usando fuzzers disponíveis. O estudo envolve o aprendizado de formatos de representação de imagens, como JPEG e PNG, por exemplo. O projeto envolve também o aprendizado sobre fuzzing, e o uso de ferramentas de apoio. O projeto pode ser realizado por 2 alunos, cada qual responsável por aplicar uma técnica diferente de fuzzing.	Em Equipe
	Uso de testes baseados em modelos em combinação com o desenvolvimento dirigido pelo comportamento	O desenvolvimento dirigido pelo comportamento, mais conhecido como BDD (Behavior-Driven Development) é um paradigma usado em ambiente ágil que usa uma linguagem ágil, denominada Gherkin. A linguagem é utilizada pois pode ser compreendida por diversos participantes: clientes, desenvolvedores, equipes de teste. A linguagem Gherkin é usada para descrever requisitos na forma de cenários, os quais são transformados em casos de teste de Aceitação do sistema. Estes cenários são criados manualmente pelas equipes de desenvolvimento ou de testes. Uma vantagem é que a qualidade dos testes vai depender da experiência de quem os escreve. Uma outra desvantagem é que a equipe de testes não dispõe de uma métrica de cobertura que permita determinar se os testes são suficientes ou não. A combinação de BDD com testes baseados em modelos ou MBT (do inglês Model-Based Testing) pode ser uma solução para esses problemas. Em MBT, os casos de teste são produzidos a partir de um modelo representando o comportamento do sistema. O objetivo deste projeto é estabelecer um método que permita às equipes de teste, a partir dos requisitos modelados na forma de modelos de estados da UML, derivarem os casos de teste em Gherkin. No caso de mudança nos requisitos, por exemplo, com o acréscimo de um novo incremento, os casos de teste necessários são produzidos automaticamente a partir do modelo atualizado. Para que o método possa ser usado em ambientes de desenvolvimento ágil, o apoio de ferramentas é fundamental. Neste projeto o aluno deve entender, instalar e utilizar as ferramentas de apoio. Como estudo de caso, uma aplicação Web deve ser utilizada.	Individual
Flávio Keidi Miyazawa	Problemas de Corte e Empacotamento.	Os alunos deverão estudar algoritmos para problemas de empacotamento bidimensional em placas. Alguns alunos poderão investigar a versão de empacotamento em placas, outros em faixa, versões unidimensional, bidimensional, tridimensional ou multidimensional. Versões poderão ser de empacotamentos de itens retangulares, caixas ou mesmo irregulares. Direções de estudo poderá ser prática implementando e comparando os algoritmos mais promissores ou teórica, com estudo e desenvolvimento de algoritmos de aproximação, ou nas duas linhas. Os projetos poderão ser individuais ou em equipe.	Individual ou em Equipe
	Problemas de Roteamento de Veículos.	Os alunos deverão estudar algoritmos para problemas de roteamento de veículos. Cada veículo deve percorrer uma rota, partindo de um depósito, atender clientes e voltar novamente ao depósito. Cada cliente deve ser atendido por um veículo. O objetivo é minimizar uma função de custo associada às rotas. Cada projeto focará em uma versão do problema considerando características mais específicas. Os projetos poderão ser individuais ou em equipe.	Individual ou em Equipe
Gerberth Adin Ramirez Rivera	Processamento de imagens com redes neurais baseadas em grafos	As redes neurais baseadas em grafos são uma opção interessante para processar imagens já que usam relações mais complexas para o processamento. O projeto visa o estudo destas técnicas para tarefas em imagens (classificação, segmentação, geração, entre outras).	Individual
	Aprendizado não supervisionado probabilístico de representações	O aprendizado não supervisionado é um problema importante a ser resolvido já que a criação de etiquetas é um processo caro e complexo. Um dos paradigmas usados para aprender representações que são usadas em outras tarefas posteriores é contrastar representações para obter uma distribuição baseada na energia entre os dados. Este projeto visa aprender outras distribuições (paramétricas ou não) para melhorar as tarefas posteriores de aprendizado.	Individual
Julio Cesar dos Reis	Alinhamento de grafos de conhecimento	Grafos de conhecimento definem fatos expressos como triplas considerando sujeito, predicado e objeto na representação do conhecimento. Usualmente diversos grafos de conhecimento são publicados em um determinado domínio. É relevante criar alinhamentos tanto de classes que modelam conceitos quanto entre instâncias dessas classes definidas em diferentes grafos de conhecimento. O objetivo deste projeto é estudar técnicas de alinhamento de entidades expressas em grafos de conhecimento. Usaremos conjunto de dados existentes na Ontology Alignment Evaluation Initiative para avaliar os métodos concebidos em análises experimentais.	Individual
	Descrição semântica de publicações científicas.	Publicações científicas podem ser melhor recuperadas e analisadas quando o significado dos atributos que caracterizam os dados da publicação são codificados em modelos computacionais que representam explicitamente a semântica. O objetivo deste projeto é propor e desenvolver um sistema que coleta dados sobre artigos científicos e enriquece semanticamente os registros por meio de vocabulários que descrevem precisamente os conceitos do domínio. O trabalho envolverá estudar linguagens para a criação e consulta de ontologias.	Individual ou em Equipe
	Sistema para explorar dados interconectados abertos.	Um número crescente de dados interconectados abertos (Linked Open Data) são publicados e disponíveis em repositórios na Web. Há diversas oportunidades no uso e integração desses dados interconectados, com semântica interpretável pela máquina, em diferentes domínios. Este projeto visa construir funcionalidades de software para consultar e combinar fatos descritos nestes repositórios. O trabalho exigirá o estudo de uma linguagem de consulta para acesso a fontes de dados na Web Semântica (SPARQL).	Individual
	Sistemas de questão e respostas usando bases RDF.	Sistemas de questões e respostas fazem parte de um esforço contínuo para aprimorar a interação homem-computador. Este projeto objetiva implementar um sistema que permita interpretar uma questão em linguagem natural e obter uma consulta estruturada. Visamos considerar consultas em grafos de conhecimento descritos em RDF. O trabalho exigirá o estudo de uma linguagem de consulta para acesso a fontes de dados na Web Semântica (linguagem SPARQL). As respostas obtidas serão convertidas em uma representação final para o usuário.	Individual
	Visualização de ontologias.	Ontologias permitem representar conceitos em um domínio e podem ser úteis para usuários fazerem sentido de conceitos e suas relações. Contudo, poucos estudos investigam a interação com essas estruturas. Este projeto visa projetar e construir um sistema com uso de ontologias para permitir usuários navegarem entre conceitos de disciplinas do curso de engenharia e ciência da computação. Utilizaremos design centrado no usuário e técnicas participativas para elaborar a estrutura de visualização das ontologias. Este sistema poderá permitir que aluno(s) melhor entendam os conceitos e suas relações nas disciplinas que compõem o curso.	Individual
Luiz Fernando Bittencourt	Gerência de recursos em sistemas distribuídos	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.	Individual ou em Equipe
Nelson Luis Saldanha da Fonseca	Aprendizado Federativo para Classificação de Imagens e/ou Vídeos	Neste trabalho o aluno desenvolverá experimentos via emulação e/ou simulação de aprendizado federativo (do inglês Federated Learning) utilizando diferentes técnicas de particionamento de dados, explorando desde a visualização dos dados até a avaliação dos resultados dos algoritmos de aprendizado de máquina.	Individual
	Aprendizado Federativo para Medicina	Neste trabalho o aluno desenvolverá experimentos via emulação e/ou simulação de aprendizado federativo vertical (do inglês Vertical Federated Learning) utilizando dados de saúde, será trabalhado desde a visualização dos dados até a avaliação dos resultados dos algoritmos de aprendizado de máquina supervisionado.	Individual
	Avaliação de Desempenho de Aprendizado Federativo na Borda da Rede	Neste trabalho o aluno irá desenvolver experimentos via emulação ou simulação para avaliar diferentes aspectos relacionados à execução de aprendizado federativo (do inglês Federated Learning) em redes de computação de rede.	Individual
	Agregação Assíncrona Federativa de Modelos de Aprendizado Supervisionados	Neste trabalho o aluno desenvolverá um algoritmo de agregação assíncrona para aprendizado federativo (do inglês Federated Learning) que será integrado ao projeto FATE (Federated AI Technology Enabler). Além disso, serão desenvolvidos experimentos para comparar a agregação síncrona com a agregação assíncrona.	Individual
	Aprendizado Federativo Descentralizado	Neste trabalho o aluno desenvolverá algoritmos que permitam a execução de aprendizado federativo (do inglês Federated Learning) de forma distribuída, ou seja, sem um servidor central. Para validar os resultados, o aluno desenvolverá experimentos via emulação e simulação.	Individual
	Avaliação de desempenho do HTTP 3.0	Neste trabalho, o aluno realizará experimentos via simulação, emulação ou modelagem analítica do desempenho da nova proposta de padronização do protocolo HTTP 3.0 e protocolos de transporte usados pelo HTTP 3.0.	Individual ou em Equipe
Zanoni Dias	Diagnóstico de COVID-19 usando Redes Neurais Profundas	A pandemia COVID-19 impactou todas as esferas de nossa sociedade. A epidemia aumentou a pressão sobre os sistemas de saúde pública, incluindo a comunidade científica a desenvolver e avaliar métodos para diagnosticar pacientes de forma confiável. Impulsionadas por sua eficácia na análise de imagens médicas, as redes neurais profundas têm sido vistas como uma possível alternativa para diagnosticar automaticamente pacientes com COVID-19 a partir de radiografias de tórax. O objetivo deste trabalho é avançar nos resultados previamente obtidos no artigo "COVID-19 X-ray Image Diagnostic with Deep Neural Networks", vencedor do Best Paper Award do Brazilian Symposium on Bioinformatics (BSB'2020).	Individual
	Identificação de Estruturas Secundárias e de Funções de Proteínas usando Transformers	Transformers são modelos de aprendizado de máquina usados principalmente em Processamento de Linguagem Natural (NLP). Recentemente estes modelos têm sido empregados com sucesso em outros domínios, como imagens. O objetivo deste projeto é empregar transformers para identificação de estruturas secundárias e de funções de proteínas, já que seqüências de aminoácidos (proteínas) podem ser interpretadas como (longas) frases, onde cada aminoácido pode ser tratado como uma palavra.	Individual
	Classificação de Imagens de Proteínas usando Redes Neurais Profundas	Com os avanços da tecnologia, principalmente em relação à microscopia de alta resolução, imagens de proteínas são geradas em uma velocidade maior do que se pode avaliar manualmente. Esse tipo de imagem é normalmente usada em pesquisas biomédicas que podem ser a chave para o próximo avanço na biotecnologia. Este trabalho tem como objetivo empregar redes neurais convolucionais profundas para automatizar a análise dessas imagens.	Individual
	Rearranjo de Genomas com Genes Multiplicados	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 nos genomas capazes de alterar a seqüência genética. Quando os genomas apresentam genes multiplicados, eles são representados por strings. O objetivo deste projeto é o desenvolvimento de heurísticas para problemas de rearranjo de genomas com estas características.	Individual
	Diagnóstico de Patologias em Imagens de Raios-X usando Redes Neurais Profundas	Muitas patologias são diagnosticadas por especialistas através de imagens de raios-X. O objetivo deste trabalho é explorar diversos problemas de identificação de patologias em bases públicas de imagens de raios-X, de forma a propor uma metodologia baseada em redes neurais profundas que possa aproveitar o conhecimento obtido numa base (para uma ou mais patologias específicas) em outra(s) base(s) (com um conjunto distinto de patologias).	Individual
	Aumento de Dados para Classificação de Imagens Utilizando Redes Neurais Profundas	Pesquisas em Aprendizado de Máquina Profundo têm trazido grandes transformações para diversos setores da sociedade. Apesar disso, a escassez de dados em determinados domínios de aplicação ainda representa um dos principais problemas enfrentados por esse campo de pesquisas, desencadeando desafios relacionados ao desempenho e à justiça dos algoritmos. O objetivo deste projeto é investigar técnicas de aumento de dados para a adição de volume e diversidade em conjuntos de dados destinados a tarefas de classificação de imagens utilizando Aprendizado de Máquina Profundo.	Individual
	Usando Aprendizado por Reforço para Computar Distância de Rearranjo de Genomas	O objetivo deste projeto é usar técnicas de Aprendizado de Máquina, mais especificamente de Aprendizado por Reforço, para construir um modelo capaz de computar a distância de rearranjo entre genomas (representados por permutações). Para este projeto iremos considerar as versões do problema de Rearranjo de Genoma onde são permitidos os eventos de reversão e transposição.	Individual
Ordenação de Permutações por Operações Determinísticas	Problemas de Rearranjo de Genomas podem ser tratados como problemas de ordenação de permutações. Tipicamente, eventos de rearranjo podem movimentar elementos para qualquer posição da permutação. Neste projeto pretende-se explorar variações onde as operações são determinísticas, ou seja, os elementos devem ser movidos para posições específicas da permutação. As operações determinísticas a serem exploradas neste projeto são: inserções, trocas de pares e reversões.	Individual	
Bioinformática para Análise de Transcriptomas	Transcriptoma refere-se ao conjunto completo de transcritos (RNAs mensageiros, RNAs ribossômicos, RNAs transportadores e os microRNAs) de um dado organismo e é o reflexo direto da expressão dos genes. Uma das técnicas mais populares de geração de dados em genômica é o sequenciamento de RNA (RNAseq), devido ao seu relativo baixo custo. O objetivo deste projeto é desenvolver metodologias para associar os RNA sequenciados aos genes do organismo.	Individual	