

Professor Responsável	Título	Descrição dos Projetos Finais de Graduação para o 2s2020	Individual/Equipe
Alexandre Xavier Falcão	Avaliação de modelos de aprendizagem profunda para segmentação semântica de imagens aéreas.	Nesse projeto, o aluno terá a oportunidade de estudar segmentação semântica de imagens com aprendizado profundo. O Objetivo do projeto é avaliar modelos propostos para o problema de segmentação semântica de imagens aéreas. O estudo envolve avaliar a capacidade de generalização dos modelos para um novo conjunto de imagens diferente do utilizado para o treinamento do modelo. O estudo envolve a utilização de pacotes em Python para implementação e avaliação dos modelos.	Individual
	Avaliação de técnicas de projeção não-linear de dados multidimensionais em aprendizado semi-supervisionado de máquina	Neste projeto, o aluno terá a oportunidade de estudar técnicas de projeção não-linear de dados multidimensionais e técnicas de aprendizado semi-supervisionado de máquina. O objetivo do projeto é avaliar o impacto de algumas técnicas de projeção não-linear na anotação automática de imagens. O estudo envolve técnicas como LAMP, UMAP, t-SNE, e redes neurais, usando pacotes disponíveis em Python.	Individual
Breno Bernard Nicolau de França	Aplicativo Mapa Afetivo	Este projeto deve ser executado em colaboração com um grupo de pesquisa em Artes da Cena, representada pela pesquisadora Marília Ennes Becker. O aplicativo apresenta o mapa da Unicamp para o usuário registrar pontos específicos de seu percurso que lhe remetam a uma experiência singular e subjetiva vivida em determinado lugar. A proposta é incentivar práticas sensíveis de ocupação do espaço que tenham caráter lúdico e afetivo. Um convite para subverter a lógica funcional de ocupação objetiva do espaço e permitir-se acessar novos trajetos lançando um olhar mais atento e curioso para o percurso e suas pausas. Um mapa interativo e colaborativo que apontará paradas e portanto, promoverá percursos que potencializam a experiência de troca sensível entre o corpo que atravessa ou pausa e o espaço que acolhe. Por entre os caminhos possíveis pode haver uma sombra de árvore que seja boa para contar histórias, um vão entre os prédios que é perfeito para soar uma nota sob o raio de sol no fim da tarde, ou ainda, o pilar direito da marquise que esconde um chão ideal para dançar a dois; as possibilidades são infinitas, por entre as ruas, esquinas, praças, árvores, institutos da Unicamp e o simples ato de caminhar. Tecnologias ainda em aberto. Pode ser em equipe de até 3 pessoas.	em Equipe
	Análise de repositórios para entrega contínua	Extensão da ferramenta CDM para análise de repositórios de software livre que utilizam práticas de entrega contínua	Individual
	Qualitative Data Analytics	Desenvolvimento de ferramenta para análise qualitativa de dados.	Individual
Eliane Martins	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 ubíqua, denominada Gherkin. A linguagem é ubíqua 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 desvantagem 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 permite determinar se os testes que escreve 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
	Teste de segurança de aplicações Java usando fuzzers	Fuzz testing (ou fuzzing) é uma técnica de teste de software que insere dados inválidos ou aleatórios, também chamados de fuzz, no sistema de software para descobrir erros funcionais, bem como brechas de segurança. Os dados são inseridos usando técnicas de teste automatizadas ou semi-automáticas, após as quais o sistema é monitorado para determinar se ocorreram exceções, queda (crash) ou bloqueio do sistema. Estudo das brechas de segurança mais comuns em Java, e também da técnica utilizada (fuzzing), são parte inicial do trabalho. O objetivo deste trabalho é utilizar alguns fuzzers para Java para analisar sua capacidade de determinar brechas de segurança.	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
Hélio Pedrini	Reconhecimento de Ações em Vídeo.	Identificação de ações humanas em seqüências de vídeos por meio de técnicas de processamento de imagens, análise de vídeos, visão computacional e aprendizado de máquina.	Individual
	Identificação de Eventos Anômalos em Vídeos de Vigilância.	Identificação de comportamento anômalo em vídeos de vigilância baseada em técnicas de visão computacional, análise de imagens e aprendizado de máquina.	Individual
	Análise de Imagens e Vídeos.	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).	Individual
Juliana Freitag	Smart Campus	O projeto Smart Campus Unicamp (https://smarcampus.prefeitura.unicamp.br/) desenvolve soluções baseadas em Internet das Coisas com o objetivo de melhorar a qualidade de vida no campus e fazer uso mais eficiente e sustentável dos recursos. Esse projeto visa desenvolver uma solução baseada em Internet das Coisas para o campus. A solução a ser desenvolvida será definida em comum acordo entre aluno e professora.	Individual ou Em Equipe (discriminar na descrição a divisão dos trabalhos)
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, tais como formas de implementação da aplicação, análise de desempenho e algoritmos de alocação de recursos.	Individual ou em Equipe
Nelson Luis Saldanha da Fonseca	Aprendizado Federativo Vertical em Computação de Borda	O objetivo deste trabalho é desenvolvimento de funcionalidades que permitam Vertical Federated Learning (VFL) na borda da rede. Neste modelo o aluno deverá desenvolver um ambiente que contenha dispositivos móveis e dispositivos IoT. VFL é capaz de tirar vantagem da heterogeneidade de características de banco de dados distribuídos mantidos por empresas a fim de criar modelos de aprendizado de máquina otimizados sem expor dados privados. O modelo VFL é adequado para aplicações e serviços Business-to-Business (B2B). Um exemplo de ambiente inclui hospitais e farmácias, onde estes possuem clientes (ou pacientes) em comum, aqui os hospitais podem contribuir com as farmácias usando prontuários médicos para contribuir no tratamento de doenças e reduzir hospitalizações, portanto existem objetivos diferentes, mas pacientes em comum.	Individual ou em Equipe
	O paradigma de Multi-Task Learning aplicado em Computação de Borda Federativa	O objetivo deste trabalho é desenvolver um ambiente de computação de borda capaz de suportar meta-task learning. O multi-task learning será utilizado para permitir que diferentes tarefas sejam executadas ao mesmo tempo. Aplicando multi-task learning de forma efetiva é possível identificar pontos semelhantes entre os dados dos usuários e assim agrupá-los, ou seja, um grupo de dispositivos de borda forma uma tarefa. Esta formação de tarefas deverá ser definida antes do deployment de modelos de aprendizado de máquina. Usando deste mecanismo, é possível reduzir problemas de comunicação, e tolerância à falha.	Individual ou em Equipe
	Serveless em Redes de Computação de Borda	O objetivo deste trabalho é inserir o paradigma de serveless na borda da rede. O mecanismo a ser desenvolvido irá permitir o uso do conceito de funções de borda. A ideia do serverless é que um servidor não fique trabalhando dedicado a um serviço/aplicação somente, assim o conceito de funções surge. Uma função pode ser, por exemplo o carrinho de um ecommerce. Aqui as aplicações devem ser divididas estrategicamente em funções. Um dos benefícios de usar serverless está na capacidade de redução de custos.	Individual ou em Equipe
	Política justa para treinamento de modelos federativos na borda da rede	Neste trabalho o objetivo é explorar um dos desafios de Federated Learning (FL) em computação de borda é trabalhar com dados distribuídos não identicamente e não independentemente. Como possível solução será desenvolvido um algoritmo justo dentre os dispositivos de borda. Este algoritmo será capaz de limitar a contribuição do dado dos dispositivos móveis que receberão modelos de aprendizado de máquina.	Individual ou em Equipe
	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

Professor Responsável	Título	Descrição dos Projetos Finais de Graduação para o 2s2020	Individual/Equipe
Zanoni Dias	Bioinformática Pragmática na Análise de Transcriptomas.	Uma das técnicas mais populares de geração de dados em genômica é o sequenciamento de RNA (RNAseq), devido ao seu relativo baixo custo. Quando o organismo com o qual se trabalha já possui o genoma descrito, os desafios computacionais para a análise estão razoavelmente definidos. Quando o genoma de referência não existe, entretanto, o problema admite soluções distintas e incompatíveis entre si, com grande taxa de falsos positivos. O objetivo deste projeto é desenvolver metodologias que gerem listas de genes com alta proporção de verdadeiros positivos.	Individual
	Classificação de Anormalidades Musculoesqueléticas	Este projeto tem como objetivo investigar técnicas de classificação de imagens radiográficas para detectar anormalidades musculoesqueléticas, fazendo uso de modelos de Aprendizado de Máquina. A base de dados utilizada neste projeto será a base "MURA: Large Dataset for Abnormality Detection in Musculoskeletal Radiographs" (https://arxiv.org/abs/1712.06957), que consiste de um grande conjunto de imagens de Raios-X ósseos.	Individual
	Classificação de Problemas Cardiorpulmonares	Este projeto tem como objetivo investigar técnicas de classificação de imagens radiográficas para detectar problemas cardiorpulmonares, fazendo uso de modelos de Aprendizado de Máquina. A base de dados a ser utilizada neste projeto é denominada "CheXpert: A Large Chest Radiograph Dataset with Uncertainty Labels and Expert Comparison" (https://arxiv.org/abs/1901.07031), que consiste de um grande conjunto de imagens de Raios-X de tórax.	Individual
	Classificação de Estruturas Secundárias de Proteínas	O objetivo deste trabalho é desenvolver uma abordagem capaz de prever estruturas secundárias de proteínas, usando técnicas de Aprendizado de Máquina. Neste projeto, utilizaremos a base de dados pública PDB (http://www.rcsb.org), que contém mais de 150 mil proteínas classificadas.	Individual
	Classificação de Funções de Proteínas	O objetivo deste trabalho é desenvolver uma abordagem capaz de prever funções de proteínas, usando técnicas de Aprendizado de Máquina. Neste projeto, utilizaremos a base de dados pública PDB (http://www.rcsb.org), que contém mais de 150 mil proteínas classificadas.	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 projetos são: inserções, trocas de pares e reversões.	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 programa 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 os eventos de reversão e transposição são permitidos.	Individual