Professor Responsável	Título	Descrição dos Projetos Finais de Graduação para o 2s2022	Individual/Equipe
Julio César López Hernández	**Implementação em Software de Criptografia Leve**	Implementação em Software de Algoritmos Simétricos Um algoritmo simétrico (ou de chave secreta) é um método para encriptar/desencriptar mensagens entre duas entidades que compartilham uma chave (um segredo). No processo de encriptação, o algoritmo simétrico junto com a chave transforma uma mensagem numa cadeia de bits que não pode ser entendida por ninguém que não possua a mesma chave utilizada para encripta-la. Uma vez que o destinatário da mensagem, quem possui a mesma chave, recebe a mensagem cifrada, o algoritmo simétrico junto com a chave reverte a operação, recuperando o texto da mensagem original. Dessa forma, os algoritmos simétricos são utilizados para comunicações seguras entres duas entidades (ou pessoas) que compartilham uma chave secreta. Em 2001, o Instituto Nacional de Padrões Americano (NIST) anunciou o novo padrão da criptografia simétrica, chamado AES (Advanced Encryption Standard). É um algoritmo que utiliza chaves de 128, 192 ou 256 bits para cifrar mensagens de 128 bits. Embora, o AES seja suportado em muitas plataformas, com instruções em hardware para acelerar o processo de cifragem, o AES não foi projetado para aplicações em dispositivos limitados. Recentemente, o NIST anunciou uma lista de 10 algoritmos candidatos para criptografia leve, isto é, algoritmos simétricos que foram projetados para dispositivos limitados. O objetivo principal deste projeto é estudar algoritmos criptográficos simétricos, no contexto da Criptografia Leve, e desenvolver códigos otimizados em C para diferentes plataformas arm.	**Individual**
		As principais atividades do projeto são: 1. Estudar os conceitos básicos de criptografia simétrica (encriptação, funções de resumo, ver referências [1] e [2]) 2. Estudar 2 algoritmos da lista: ASCON, Elephant, GIFT-COFB, Grain-128AEAD, ISAP, Xoodyak, TinyJambu, SPARKLE, Romulus, PHOTON-Beetle. A especificação de cada algoritmo é dada em referência [3]. 3. Estudar técnicas de programação para armv8/armv9. 4. Desenvolver implementações em C dos algoritmos selecionados. 5. Estudar ataques de canal lateral (side-channel attacks) em software (implementações em tempo constante). 5. Desenvolver implementações otimizadas em C dos algoritmos selecionados. 6. Realizar medições de tempo e comparar com outras bibliotecas. 7. Escrever um relatório, e se possível submeter um artigo para publicação. Requisitos: Motivação para aprender novos algoritmos e matemáticas, bons conhecimentos da linguagem C.	
		Referências: [1]. Cryptography and Network Security, William Stallings, 7 ed, 2020 [2]. Cryptography: Theory and Practice, 4 ed. Stinson, Paterson, 2019 [3]. https://csrc.nist.gov/Projects/lightweight-cryptography/news	
Edmundo Roberto Mauro Madeira	**Aprendizado de Máquina aplicado a alocação de recursos de rede**	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.	**Individual**
	Aprendizado de Máquina aplicado para predição de mobilidade de usuários	Aplicações e serviços hospedados em uma infraestrutura de rede podem sofrer variações de demanda ao longo do tempo devido a diferentes fatores. Em especial, a mobilidade dos usuários trás uma série de fatores que contribuem para a dinamicidade deste cenários. Entretanto, vários estudos têm demonstrado que há certa previsibilidade na movimentação desses usuários. Este fator pode ser explorado para melhorar o gerenciamento da rede. Este projeto tem como objetivo estudar soluções baseadas em aprendizado de máquina para otimizar o gerenciamento da rede a partir da predição da mobilidade dos seus usuários.	**Individual**
Eliane Martins	**Uso de testes baseados em modelos em combinação com o desenvolvimento dirigido pelo comportamento**	Em projetos prévios foi utilizado o modelo de estados para criação automática de cenários de testes em Gherkin, linguagem ubíqua utilizada no desenvolvimento dirigido pelo comportamento, mais conhecido pela sigla BDD (Behavior Driven Development). Vimos que os cenários produzidos pelo modelo são maiores e exercitam mais funcionalidades do sistema do que os cenários produzidos manualmente. Nosso objetivo agora é determinar a efetividade destes cenários produzidos automaticamente na detecção de defeitos. Para isso, utilizaremos uma técnica denominada Análise de Mutantes, que consiste em introduzir defeitos no código para analisar a capacidade de conjuntos de teste em revelar a presença destes defeitos. Uma aplicação Web será utilizada como estudo de caso. Os cenários produzidos manualmente para essa aplicação serão comparados aos cenários produzidos a partir do modelo de estados.	**Individual**
	Uso de gêmeos digital nos testes de sistemas ciber-físicos	Um gêmeo digital (digital twin) é uma réplica virtual de um objeto ou sistema físico, podendo receber como entrada dados em tempo real de objetos ou sistemas físicos e, com simulações, predizer a saída produzida por estes objetos ou sistemas em resposta às entradas recebidas. Com isso é possível ter ideias sobre o desempenho ou problemas que possam surgir durante a execução.O objetivo do projeto é o estudo dessa tecnologia, as plataformas de apoio existentes e utilização de uma destas plataformas em um exemplo de um sistema ciber-físico. O projeto pode ser feito em equipe, em que cada um será responsável por criar um gêmeo digital para um dispositivo utilizado pelo sistema ciber-físico escolhido.	**Equipe**
Hélio Pedrini	**Reconhecimento de Ações em Vídeo.**	Identificação de acões humanas em sequências 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, bologia, biometria, microscopia, vigilância e segurança).	**Individual**
Hervé Cédric Yviquel	**Biblioteca de álgebra linear para supercomputadores**	O objetivo desse projeto é implementar uma biblioteca de álgebra linear para facilitar computação em matriz esparsa para supercomputadores (ou clusters). Essa biblioteca deve ser implementada em C/C++ em cima do runtime de tarefa distribuída do projeto OmpCluster usando o modelo de programação OpenMP. Se for em equipe, cada um pode implementar operações diferentes: por exemplo multiplicação matriz-vetor, multiplicação matriz-matriz, etc.	**Individual ou em Equipe**
	Aplicar o modelo de desempenho Roofline para aplicações científica baseadas em tarefa para supercomputadores	O objetivo deste projeto é melhorar a análise de desempenho de aplicação científica baseado em tarefas usando o modelo roofline. O modelo roofline foi desenvolvido pelo Berkeley Lab para visualizar facilmente os gargalos e potenciais de desempenho dos kernels de computação e pode ser facilmente usado usando ferramentas como Timemory. O estudante integrará a ferramenta Timemory no runtime de tarefa OmpCluster e explorará a melhor maneira de apresentar o perfilamento ao usuário.	**Individual**
Julio Cesar dos Reis	**Interfaces vestíveis para a captura e rotulação de dados fisiológicos**	A captura, processamento e interpretação de dados fisiológicos podem desenvolver um papel relevante no comportamento de sistemas interativos. Este projeto visa explorar artefatos existentes no mercado que permitam a coleta de dados fisiológicos, como smartwatches. Visamos projetar e construir uma aplicação que permita capturar esses dados a partir de instrumentos existentes e desenvolver cenários adequados para a rotulação desses conforme aspectos emocionais das pessoas. Objetivamos gerar e analisar um conjunto de dados que possa ser útil para detectar as emoções das pessoas a partir de sua frequência cardíaca e outros sinais fisiológicos.	**Individual**
	Avaliação na plataforma OpenDesign	O Opendesign é uma plataforma online para o apoio na condução colaborativa e distribuída do design de sistemas computacionais. A condução de processos de avaliação do design desenvolve um papel chave para o aprimoramento de features no software e para o informar o redesign em estudo de um projeto. A plataforma OpenDesign demanda a construção de novos mecanismos que suportem designers e outros stakeholders envolvidos na condução e documentação de avaliação. O objetivo deste projeto é conceber, implementar e avaliar ferramentas para o registro de diferentes tipos de métodos de avaliação na plataforma.	**Individual ou em Equipe**
	Processamento de diálogos textuais na língua Portuguesa	Sistemas de chatbots desempenham papel chave como uma ferramenta de autoatendimento a clientes em grandes empresas. A construção desses sistemas ainda é repleta de desafios no design da experiência do usuário e no funcionamento do sistema visando uma experiência agradável de uso. Por exemplo, clientes conseguirem resolver problemas menos complexos sem necessitar de diversos atendimentos ou delongadas interações com atendentes humanos. Sistemas de chatbots requerem uso de técnicas de processamento de linguagem natural. Este projeto visa investigar o reconhecimento de entidades no processamento de texto na língua Portuguesa em sistemas conversacionais computadorizados. Exploraremos modelos de classificação automática de entidades.	**Individual**
	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 à 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 à 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 à 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**

Professor Responsável	Título	Descrição dos Projetos Finais de Graduação para o 2s2022	Individual/Equipe
Luiz Fernando Bittencourt	**Aprendizado de máquina federado no simulador MobFogSim**	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.	**Individual ou em Equipe**
	Elasticidade e auto-distribuição	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.	**Individual ou em Equipe**
	Gerência de Sistemas Auto-adaptativos	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.	**Individual ou em Equipe**
	Aprendizado de Máquina Distribuído	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.	**Individual ou em Equipe**
	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**
	Monitoramento de colmeias com Internet das Coisas	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).	**Individual ou em Equipe**
Marcos Medeiros Raimundo	**Biblioteca de otimização multi-objetivo para aprendizado de máquina**	O uso de otimização multi-objetivo em aprendizado de máquina tem se mostrado capaz de lidar com multiplos desafios, incluindo geração de ensemble, classes desbalanceadas, interpretabilidade e imparcialidade. Para tanto, é necessário que o usuário seja capaz de discriminar objetivos conflitantes no modelo de aprendizado de máquina, um dos evidentes é o conflito entre a complexidade do modelo e erro de aprendizado durante o treinamento. Após essa modelagem, é gerado uma sequência de modelos de consistem diferentes compromisos entre complexidade e erro de aprendizado, que então podem ser usados numa tomada de decisão. Este projeto visa criar um software que é similar e mistura elementos do sklearn, tensorflow e scipy/guroby para facilitar o uso de otimização multi-objetivo em aprendizado de máquina.	**Individual ou em Equipe**
	Otimização multi-objetivo em redes neurais	O uso de otimização multi-objetivo em aprendizado de máquina tem se mostrado capaz de lidar com multiplos desafios, incluindo geração de ensemble, classes desbalanceadas, interpretabilidade e imparcialidade. Para tanto, é necessário que o usuário seja capaz de discriminar objetivos conflitantes no modelo de aprendizado de máquina, um dos evidentes é o conflito entre a complexidade do modelo e erro de aprendizado durante o treinamento. O grande desafio de modelos de otimização multi-objetivo em redes neurais (e outros modelos de maios complexidade) reside na não convexidade resultante do problema. Este projeto visa aprimorar um algoritmo de otimização multi-objetivo para o uso em modelos mais complexos de aprendizado de máquina.	**Individual**
Zanoni Dias	**Visualização e interpretação de modelos profundos**	A chegada da aprendizagem profunda, bem como seu emprego nos mais diversos setores de nossa sociedade, provocou um crescente interesse em encontrar-se técnicas e metodologias confiáveis de visualização e interpretação de resultados, a fim de se comprovar o funcionamento correto e não-enviesado do modelo de decisão. Este projeto tem como objetivo estudar as diferentes técnicas de visualização de resultados existentes, focando-se em baseadas em gradiente, Class Activation Mapping (CAM) e/ou Layer-wise Relevance Propagation (LRP).	**Individual**
	Contando biodiversidade em câmeras de diferentes ecossistemas	O monitoramento da densidade de biodiversidade é uma importante atividade realizada por ecologistas em todo mundo. Ela é normalmente realizada com o posicionamento de câmeras em ambientes naturais, o registro de grandes intervalos de tempo, e a análise manual posterior. Neste sentido, a competição iWildCam 2021- FGVC8 foi proposta a fim de atrair grupos interessados em aperfeiçoar a solução para este problema. Neste projeto, técnicas de detecção de objetos e object tracking serão exploradas a fim de contar o número de animais de uma mesma espécie observados pelas câmeras.	**Individual**
	Classificação de Funções Moleculares de Proteína utilizando Processamento de Linguagem Natural	Determinar as funções que cada proteína exerce requer grande esforço técnico e monetário. Com isso, o uso de aprendizado de máquina pode auxiliar na realização desta tarefa. O objetivo deste projeto é empregar técnicas de Processamento de Linguagem Natural (NLP) para classificar funções de proteínas, com base nas suas sequências de aminoácidos, considerando que proteínas podem ser interpretadas como (longas) frases, onde cada aminoácido pode ser tratado como sendo uma palavra.	**Individual**
	Análise de Características para a Predição de Funções de Proteínas	Determinar as funções que cada proteína exerce requer grande esforço técnico e monetário. Com isso, o uso de aprendizado de máquina pode auxiliar na realização desta tarefa. O objetivo deste projeto é analisar características em que proteínas estão envolvidas para classificar as funções exercidas em diversos contextos.	**Individual**
	Ordenação Semi-Completa por Rearranjo de Genomas	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 um solução válida, que é estabelecida como entrada para o problema e considera as posições dos genes.	**Individual**
	Problemas de Partição de Strings	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 heurísticas para problemas de Partição de Strings.	**Individual**