

VII WORKSHOP EM CIÊNCIA DE DADOS

JORNADA DE CIÊNCIA DE DADOS – PROGRAMA DE VERÃO DO LNCC 2020

Auditório A, LNCC/MCTIC, Petrópolis, RJ, 14 de fevereiro de 2020

Organização: Fabio Porto (LNCC), Eduardo Ogasawara (CEFET-RJ), Artur Ziviani (LNCC)

PROGRAMAÇÃO

09:00-09:30	Abertura e Visão Geral das Atividades de Ciência de Dados no LNCC Artur Ziviani (LNCC) e Fabio Porto (LNCC)
09:30-10:30 Keynote 1	Interface Cérebro-Máquina para Reabilitação Motora pós-Acidente Vascular Cerebral via Inteligência Computacional Alex Borges Vieira (UFJF)
10:30-11:00	COFFEE BREAK
11:00-11:30 Palestra Convidada 1	Computação Afetiva: reconhecimento e simulação de emoções Gustavo Paiva Guedes e Silva (CEFET-RJ)
11:30-12:30 Sessão DEXL/LNCC	Data Science Suite Rafael P. Silva (LNCC) Aprendizado de Máquina Indutivo na Completação de Grafos de Conhecimento Incertos Daniel N. R. Silva (LNCC) Gypscie: Providing support for ML life cycle Fabio Porto (LNCC)
12:30-14:00	ALMOÇO
14:00-15:00 Keynote 2	Albert Einstein – Transformando a saúde por meio de Big Data & Analytics Andrea Suman (Sociedade Beneficente Israelita Albert Einstein)
15:00-15:30 Palestra Convidada 2	Novas formas de visualizar e analisar dados na era do big data: da indústria 4.0 à saúde 4.0 Maria Luiza Reis (Lab245 e ASSESPRO-RJ)
15:30-16:30	Breve encerramento e coffee break de <i>networking</i>

Promoção:



DEXL
 DATA EXTREME LAB

<http://dexl.lncc.br>



Agradecimentos:

Laboratório Nacional de Computação Científica (LNCC)

Av. Getúlio Vargas, 333 – Quitandinha – 25651-075 – Petrópolis, RJ

Keynote 1 (09:30-10:30)

Alex Borges Vieira (UFJF)

Título:

Interface Cérebro-Máquina para Reabilitação Motora pós-Acidente Vascular Cerebral via Inteligência Computacional

Resumo:

Apesar dos progressos recentes, o entendimento dos mecanismos e princípios que regem as a cognição e as funções complexas do cérebro ainda apresentam lacunas. Ciência de Dados, em especial, Ciência de Redes, contribui significativamente para tratar esses desafios. De fato, ao abordar o problema de forma multidisciplinar, com um forte embasamento computacional, pesquisadores são capazes de mapear, gravar, analisar e modelar elementos e interações de sistemas neurobiológicos.

Nesta palestra, inicialmente mostraremos como Ciência de Dados/Redes pode ser utilizada para ajudar a mapear funções complexas do cérebro. De forma simples, mostraremos como o cérebro pode ser mapeado em uma rede e, quais os principais problemas em analisar e tratar essa notória rede complexa.

Por fim, vamos apresentar o uso de técnicas de inteligência computacional em interface cérebro-máquina para reabilitação motora pós-acidente vascular cerebral. De forma simples, a ICM tem por objetivo prever a tarefa que uma pessoa deseja realizar, a partir de informações coletadas por sensores. Entre a coleta dos dados e a classificação da tarefa, há diversas etapas, como a limpeza dos sinais e aplicação filtros (espaciais e temporais). Em particular, vamos tratar do Filter Bank Common Spatial Patterns (FBCSP) e da Evolução Diferencial para melhorar o modelo otimizando o filtro espacial. Os resultados apresentados mostram que, quando comparado com o estado da arte, a nova proposta permite uma redução da quantidade de sensores, sem perda de eficiência na classificação das tarefas.

Biografia:

Alex Borges Vieira é professor associado do departamento de Ciência da Computação da Universidade Federal de Juiz de Fora. Vieira é pesquisador nível 2 do CNPq. Atualmente, atua como pesquisador visitante no Laboratório Nacional de Computação Científica (LNCC). Seus interesses de pesquisa incluem redes de sensores, caracterização, modelagem e análise de redes de computadores e ciência de redes.

Keynote 2 (14:00-15:00)

Andrea Suman (Sociedade Beneficente Israelita Albert Einstein)

Título:

Albert Einstein – Transformando a saúde por meio de Big Data & Analytics

Resumo:

Como o uso de Big Data & Analytics tem apoiado a Sociedade Beneficente Israelita Brasileira Albert Einstein desde 2016, a impulsionar a sua missão de “levar uma gota de Einstein para cada cidadão”, e assim transformar a experiência do seu paciente, buscar melhoria na saúde populacional e garantir eficiência na gestão de recursos.

Biografia:

Andrea Suman é Superintendente de Estratégia e Inteligência de Dados da Sociedade Beneficente Israelita Albert Einstein, responsável por direcionar a estratégia e governança de dados da instituição, disseminar a cultura de dados e garantir adequação à Lei Geral de Proteção à dados.

Palestrante convidado 1 (11:00-11:30)

Gustavo Paiva Guedes e Silva (CEFET-RJ)

Título:

Computação Afetiva: reconhecimento e simulação de emoções

Resumo:

A Computação Afetiva é uma área da Ciência da Computação originada com o estudo de Rosalind Picard em 1995. É o estudo e desenvolvimento de sistemas/dispositivos capazes de reconhecer e simular o afeto humano. Um dos seus maiores propósitos é reconhecer as emoções humanas com o objetivo de fornecer uma resposta apropriada a essas emoções. Essa palestra abordará uma visão sobre a Computação Afetiva e alguns estudos que estão sendo desenvolvidos pelo Laboratório de Computação Afetiva do CEFET/RJ.

Biografia:

Gustavo é professor do Departamento de Ciência da Computação do CEFET/RJ. Possui formação em Ciência da Computação pelo Centro Universitário Plínio Leite, Letras (português-alemão) pela UFRJ, possui especialização em Gerência de Tecnologia em Computação pela UFF, mestrado em Linguística pela UFRJ e doutorado em Engenharia de Sistemas e Computação pela COPPE/UFRJ. Atualmente coordena o Laboratório de Computação Afetiva no CEFET/RJ e realiza pesquisas na área da Computação Afetiva.

Palestrante convidada 2 (15:00-15:30)

Maria Luiza Reis (Lab245 e ASSESPRO RJ)

Título:

Novas formas de visualizar e analisar dados na era do big data: da indústria 4.0 à saúde 4.0

Resumo:

Desde que começamos a capturar grande volume de dados de fontes diversas não padronizadas, enfrentamos grandes desafios para tratar essa informação. O primeiro desafio é a normalização de campos, de termos e da taxonomia. O segundo desafio é criar indicadores ideais para melhor representar uma tendência ou um compilado de informações numéricas. Para melhorar a percepção da informação com indicadores, a evolução temporal deve ser apresentada e por fim, o conteúdo do dado não estruturado que é a fonte final da informação. Nesta palestra serão abordados os seguintes temas: unificação de dados e criação de indicadores para fonte de dados variados usando machine learning, a representação gráfica multidimensional de dados não numéricos e aplicações na indústria e na saúde 4.0.

Biografia:

- Presidente de ASSESPRO RJ;
- CEO da Lab245 Software;
- Pesquisadora Visitante em Visão Artificial na Université Blaise Pascal na França;
- Doutora em Engenharia Nuclear pela COPPE/UFRJ;
- Mestre em Engenharia Mecânica pela COPPE/UFRJ;
- Engenheira Mecânica/Nuclear UFRJ;
- Atualmente é pesquisadora, empresária e palestrante em congressos no Brasil e exterior sobre Artificial Vision, Machine Learning and Data Visualization

Sessão DEXL/LNCC (11:30-12:30)

Rafael P. Silva (LNCC)

Título: Data Science Suite

Resumo: This talk introduces DSS a suite of applications that support users in their data science tasks. we will discuss applications for data visualization, machine and deep learning, natural language processing, Graph data analysis and time series analysis. This suite of applications runs in a service at LNCC provided by the DEXL Lab Group and can be individually installed in a local machine using docker images.

Biografia: Rafael P. Silva holds a BSc degree in Physics by UFF (2018). Currently, he is a MSc candidate in Computational Modeling at LNCC. His research interests include image classification using Deep Learning, Data cleaning via statistical analysis, as well as pattern recognition on texts using text mining.

Daniel N. R. Silva (LNCC)

Título: Aprendizado de Máquina Indutivo na Completação de Grafos de Conhecimento Incertos

Resumo: A recente e massiva produção de dados tem fomentado a demanda por mecanismos capazes de integrar e fornecer um maior nível semântico à informação. Para esse fim, parte relevante de academia e indústria tem argumentado que grafos de conhecimento seriam as estruturas mais adequadas. Contudo, alguns desafios são inerentes ao uso dessas estruturas; em particular, incompletude, incerteza e dinamicidade. Em primeiro lugar, grafos de conhecimento são incompletos, isto é, não abarcam toda a informação de interesse. Além disso, o emprego crescente de mecanismos heurísticos e de aprendizado de máquina em sua construção lhes dá um caráter incerto. Por fim, como mecanismo de integração, esses grafos devem ser flexíveis à adição de informação. Nesse contexto, diante do potencial uso de grafos de conhecimento, técnicas de aprendizado de máquina têm de forma crescente sido desenvolvidas para enfrentar o desafio de completção. Entretanto, como será apresentado, é necessário o desenvolvimento de modelos para completção que sejam mais próximos aos cenários reais dotados de incerteza e dinamicidade. Em outras palavras, modelos que sejam indutivos e capazes de lidar com a incerteza associada aos grafos de conhecimento.

Biografia: Daniel N. R. da Silva is currently a PhD candidate in Computational Modeling at LNCC. He holds a M.Sc. degree in Computational Modeling from LNCC (2017) and a B.Sc. degree in Computer Science from UNIFESO (2014). His research interests lie at the intersection among Knowledge Engineering, Machine Learning, and Network Science.

Fabio Porto (LNCC)

Título: Gypscie: Providing support for ML life cycle

Resumo: AI has become the main challenge to enable enterprise digital transformation. As a myriad of approaches appear to solve different prediction problems, learning through data has become the major principle embracing the different AI approaches. In this talk we will present Gypscie, a system being developed at the DEXL lab to support the complete ML life cycle from data pre-processing and model training to model deployment. We expect Gypscie to support building, sharing and selecting models.

Biografia: Fabio Porto is a Senior Researcher at the National Laboratory of Scientific Computing, in Brazil. He is the founder of the DEXL Laboratory, developing R&D activities in the context of scientific data analysis and management. He holds a PhD in Informatics from PUC-Rio, with sandwich at INRIA, in 2001, and a postdoc at Ecole Polytechnique Fédérale de Lausanne (EPFL). He has more than 80 research papers published in International Conferences and Scientific Journals, including VLDB, SIGMOD and ICDE. He was the General Chair of VLDB 2018 and SBBB 2015. Since 2018 he has been a member of the SBBB steering committee, and a member of SBC and ACM.