

Títulos e Abstracts do XI WPSM

Palestras Plenárias	1
Minicursos e Miniconferências	3
Palestras Sessões Temáticas	4
Redes Complexas	4
Probabilidade	5
Cronograma preliminar	6

Palestras Plenárias

Rodney Fonseca (UFBA)

Título: *Wavelet Feature Screening*

Resumo:

An initial screening of which covariates are relevant is a common practice in high-dimensional regression models. The classic feature screening selects only a subset of covariates correlated with the response variable. However, many important features might have a relevant albeit highly nonlinear relation with the response. One screening approach that handles nonlinearity is to compute the correlation between the response and nonparametric functions of each covariate. Wavelets are powerful tools for nonparametric and functional data analysis but are still seldom used in the feature screening literature. In this talk, we present a wavelet feature screening method that can be easily implemented. Theoretical and simulation results show that the proposed method can capture true covariates with high probability, even in highly nonlinear models. We also present an example with real data in a high-dimensional setting. Joint work with Pedro Morettin and Aluísio Pinheiro.

(FAPESP) Mariana Rodrigues Motta (UNICAMP)

Título: *Unsupervised Bayesian classification with correlated functional covariates*

Resumo:

We consider unsupervised classification using a latent multinomial variable to categorize a scalar response into one of the L components of a mixture model that incorporates scalar and functional covariates. This process can be viewed as a hierarchical model with three levels. The first level models the scalar response according to a mixture of parametric distributions. The second level models the mixture probabilities using a generalized linear model with functional and scalar covariates, as proposed by Garcia et al. (2024). For correlated functional covariates, the third level accounts for the correlation among curves through a mixed model. We use basis expansions to reduce dimensionality and a Bayesian approach to estimate the parameters while

providing predictions of the latent classification vector. The method is motivated by a study aiming to identify placebo responders in a clinical trial (normal mixture model).

(FAPESP) Matthieu Jonckheere (CNRS LAAS, França)

Título: *Fermat distances: old and new developments*

Resumo:

We explain the Fermat density-based estimator for weighted geodesic distances that takes into account the underlying density of the data. The consistency of the estimator is proven using tools from first passage percolation. The macroscopic distance obtained depends on a unique parameter: we discuss the choice of this parameter and the properties of the obtained distance for machine learning tasks as well as new developments for improving its practical robustness.

Daniel Takahashi (Instituto do Cérebro, UFRN)

Título: *Stochastic dynamical systems model of vocal turn-taking and its development in marmoset monkeys*

Resumo:

The system of taking turns during vocal exchanges is fundamental to the communication of several animal species, yet their developmental origins and neural mechanisms remain elusive. Marmoset monkeys readily exchange vocalization when in acoustic contact with conspecifics. Their turn-taking capacity improves during development, decreasing the amount of overlap. We developed a stochastic dynamical systems model of marmoset monkeys based on the interactions among three neural structures ("drive," "motor," and "auditory") with feedback connectivity. Fitting the model to empirical data, we found that the noise level in the auditory sensory system decreases during development, matching the timing of improvement in the capacity to avoid overlapping calls, suggesting a major role of the auditory system in early development.

(FAPESP) Keith Levin (University of Wisconsin-Madison, EUA)

Título: *Coherence-free Entrywise Estimation of Eigenvectors in Low-rank Signal-plus-noise Matrix Models*

Resumo:

Spectral methods are widely used to estimate eigenvectors of a low-rank signal matrix subject to noise. Typically, the estimation error depends on the coherence of the signal matrix. We present a method whose entrywise estimation error is independent of coherence under Gaussian noise (spiked Wigner model), achieving optimal estimation rates up to logarithmic factors. Extensions to higher ranks and non-Gaussian settings show promising results. Additionally, new metric entropy bounds for rank- r singular subspaces under the 2-to-infinity distance are derived.

(FAPESP) Carlos Cesar Trucios Maza (UNICAMP)

Título: *Forecasting risk with range based GARCH models with leverage effect*

Resumo:

Volatility models based on intraday data outperform classical GARCH approaches. We propose a range-based GARCH model that accounts for both leverage effect and extreme observations. An extensive Monte Carlo simulation and empirical applications to US stocks support the method's performance. The approach is available through a user-friendly R package, developed in joint work with Adi Pari (UNICAMP).

(FAPESP) Karthik Bharath (University of Nottingham, UK)

Título: *Spatial and temporal shapes*

Resumo:

Consider a circle on the plane deforming into an irregular wiggly closed curve; then consider several such wiggly curves scattered in a spatial domain. The first picture is a model for how, for example, a cancer cell grows over time, while the second is a snapshot of the infiltration of cancer cells within a tumour. Of interest are statistical models for quantifying how the shape of such cells, or abstractly, how the shapes of planar curves, either change over time or are correlated in space. The challenge for statistical analysis is three-fold: infinite-dimensionality of the curves; symmetries that define its shape, such as invariance to rotations and reparameterizations; and, temporal or spatial dependence between the curves. I will discuss new developments in this direction, and present results on data from real-world applications.

Minicursos e Miniconferências

Victor Coscrato (University College Cork)

Título: *Redes Neurais com PyTorch*

Resumo:

Neste minicurso, exploraremos a implementação de redes neurais usando PyTorch, desde conceitos fundamentais como tensores e diferenciação automática até a implementação de um modelo prático. Não é necessário conhecimento prévio em PyTorch, mas é recomendável familiaridade com Python e noções básicas de aprendizado de máquina.

Thiago Rodrigo Ramos (UFSCar)

Título: *Concentration of Measure in Machine Learning*

Resumo:

O fenômeno de concentração de medida é central para entender o comportamento de modelos probabilísticos de alta dimensão em aprendizado de máquina. Vamos abordar desigualdades clássicas como as de Hoeffding e McDiarmid, suas aplicações em limites de erro de generalização, seleção de modelos e análise da geometria de dados de alta dimensão, enfatizando a intuição por trás dos resultados.

Daiane de Souza Santos (ICMC-USP)

Título: *Integração do modelo de Sobrevivência com fração de cura Tempo de Promoção e algoritmos de classificação de Machine Learning*

Resumo:

Modelos de sobrevivência com fração de cura analisam dados com uma parte de indivíduos curados. Propomos modelar a incidência com algoritmos como SVM, Extreme Gradient Boosting, Random Forests, árvores de decisão e redes neurais, permitindo capturar relações complexas. Um método de estimação foi desenvolvido e simulações mostram que o modelo supera abordagens existentes, principalmente na modelagem da incidência.

Palestras Sessões Temáticas

Redes Complexas

Andressa Cerqueira (UFSCar)

Título: *Deteção de comunidades em redes com pesos*

Resumo:

Modelos de redes frequentemente exibem estruturas de comunidade. Estudamos redes ponderadas onde os pesos das arestas são gerados a partir de variáveis Gaussianas, com médias e variâncias dependentes das comunidades. Propomos um algoritmo EM rápido e tratável que recupera as comunidades com taxa de erro ótima.

(FAPESP) André Fujita (IME-USP)

Título: *A non-parametric framework to make inferences from the generation mechanisms of empirical networks*

Resumo:

Ao invés de usar características incidentais das redes, focamos na inferência sobre os mecanismos de geração, que são as verdadeiras estruturas de interesse. Apresentamos uma

estrutura estatística não paramétrica para inferência sobre esses mecanismos, essencial quando modelos paramétricos são insatisfatórios.

(FAPESP) Keith Levin

Título: *Spectral Embeddings of Correlation Networks*

Resumo:

Estudamos redes construídas a partir de séries temporais correlacionadas. Mostramos que, sob condições adequadas, a aplicação do embedding espectral adjacente a essas redes recupera os embeddings das séries verdadeiras, codificando coeficientes de Fourier dos sinais originais.

Probabilidade

(FAPESP) Pablo Groisman (UBA, Argentina)

Título: *Barycenter and k-barycenter estimation with unknown distances*

Resumo:

Analisamos a consistência dos k-médias em espaços métricos com distâncias desconhecidas, mostrando resultados de consistência sob a convergência medida-Gromov-Hausdorff. Aplicações incluem estimativas com distâncias de Fermat e Isomap, e de barycenters de medidas usando amostras.

(FAPESP) Matthieu Jonckheere (CNRS LAAS, França)

Título: *Q-learning et cookie random walks*

Resumo:

Mostramos que o Q-learning pode falhar na aprendizagem de políticas ótimas em espaços de estados infinitos, usando argumentos regenerativos baseados em passeios aleatórios "cookie".

Daniel Takahashi (Instituto do Cérebro, UFRN)

Título: *Pathwise guessing in categorical time series with unbounded alphabets*

Resumo:

Propomos um método de adivinhação não paramétrico para séries temporais categóricas, evitando a estimação explícita de probabilidades condicionais. O método apresenta uma taxa de aprendizado independente do tamanho do alfabeto e é quase ótimo, conforme mostrado por limites inferiores minimax.

Cronograma preliminar

Horário	Quarta-feira, 25/02	Quinta-feira, 26/02	Sexta-feira, 27/02
08h00 – 08h30	Abertura		
08h30 – 09h00	Credenciamento	Minicurso: Victor Coscrato	Minicurso: Victor Coscrato
09h00 – 10h00	Keith Levin (University of Wisconsin)		
10h00 – 10h30	Coffee Break	Coffee Break	Coffee Break
10h30 – 11h00	Daiane Souza (ICMC-USP)	Thiago Ramos (UFSCar)	Comunicações Orais 2
11h00 – 12h00	Mariana Mota (UNICAMP)	Matthieu Jonckheere (LAAS-CNRS)	Karthik Bharath (University of Nottingham)
12h00 – 14h00	Almoço	Almoço	Almoço
14h00 – 15h00	Rodney Fonseca (UFBA)	Daniel Takahashi (Brain Institute - UFRN)	Sessões Temáticas
15h00 – 16h00	Carlos Trucios (UNICAMP)	Comunicações Orais 1	
16h00 – 17h30	Coffee Break / Sessão de Pôsteres 1	Coffee Break / Sessão de Pôsteres 2	