

Plano de Ensino - Tópicos em Inteligência Artificial: Aprendizado de Máquina para Ciências com Regressão Simbólica - 2025.Q1

Prof. Fabrício Olivetti de França (folivetti@ufabc.edu.br)

Ementa

Aprendizado de Máquina para Ciências (ou em inglês, Scientific Machine Learning - SciML) é uma área emergente que tem por objetivo principal a busca por um modelo preditivo que seja aderente a certas características de interesse e forneça informações relevantes para o fenômeno estudado. Essa área reduz o foco do modelo criado apenas para previsão e estimula o uso do modelo como parte de uma investigação científica para a extração de conhecimento. A Regressão Simbólica é conhecida como um modelo de regressão não-linear cujas características permitem a criação de modelos interpretáveis, incorporando conhecimento de domínio, e que permitem a análise de incertezas por padrão. Esse curso tem o objetivo principal de ensinar sobre os conceitos principais de Regressão Simbólica contextualizados em SciML e como utilizar tal ferramenta para extração de conhecimento de dados diversos

Suporte e material auxiliar

- slides de aula no site: <https://folivetti.github.io/teaching/2025-summer-teaching-2>

Avaliação

A avaliação consistirá de 1 artigo com:

- Aplicação de regressão simbólica em alguma aplicação específica, ou
- Proposta de um novo algoritmo ou novos operadores para algoritmos de regressão simbólica, ou
- Criação de novas ferramentas de pré ou pós processamento de modelos de regressão simbólica

O artigo será avaliado de acordo com sua complexidade, novidade, qualidade da escrita. O artigo deverá ser submetido a um congresso ou revista da área e entregue para o professor no dia **27/04**. Sugestões de conferências:

<https://www.sbc.org.br/eventos/calendario-de-eventos/evento/102/bracis-brazilian-conference-on-intelligent-systems>

<https://csbc.sbc.org.br/2024/>

<http://www.wikicfp.com/cfp/program?id=960>

Recuperação

A resolução ConsEPE nº 182 assegura a todos os alunos de graduação com C_F igual a D ou F o direito a fazer uso de mecanismos de recuperação.

A recuperação será feita através do envio de um novo artigo com as correções apontadas.

Atendimento ao aluno

Sextas das 13hrs às 14hrs.

Conteúdo Programático

| Semana | Assunto |
|--------|---|
| 1 | Introdução e Conceitos Básicos |
| 2 | Análise de Regressão e Reg. Simbólica |
| 3 | Programação Genética e outras abordagens |
| 4 | Ferramentas de SR e Gráficos de avaliação |
| 5 | Palestras convidadas |
| 6 | Verossimilhança e distribuições |
| 7 | Otimização não-linear |
| 8 | Validação e Seleção de modelo |
| 9 | Simplificação e Integrando conhecimento |
| 10 | Extraindo informações |
| 11 | Intervalos de confiança e de predição |
| 12 | Revisão de artigos e considerações finais |

Código de honra

- Código a ser seguido durante o curso:
 - <http://professor.ufabc.edu.br/~e.francesquini/codigodehonra/>
- Colaboração é essencial, mas plágio é anti-ético

Bibliografia

Bibliografia Básica

- Gelman, Andrew, Jennifer Hill, and Aki Vehtari. Regression and other stories. Cambridge University Press, 2020.
- Harrell, Frank E. Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis. Vol. 608. New York: springer, 2001.
- Gelman, Andrew, and Jennifer Hill. Data analysis using regression and multilevel/hierarchical models. Cambridge university press, 2006.

- Bates, Douglas. “Nonlinear regression analysis and its applications.” Wiley Series in Probability and Statistics (1988).
- Nocedal, Jorge, and Stephen J. Wright, eds. Numerical optimization. New York, NY: Springer New York, 1999.
- Meeker, William Q., Gerald J. Hahn, and Luis A. Escobar. Statistical intervals: a guide for practitioners and researchers. Vol. 541. John Wiley & Sons, 2017.