

Instruções para os Relatórios da Disciplina de Aprendizado de Máquina 2024

Data de entrega: 25/06/2024

Os relatórios devem ser enviados por e-mail para: walter@azevedolab.net

Apresentação

A disciplina será avaliada de forma única por meio da entrega de um relatório com foco na modelagem de sistemas por abordagens de aprendizado de máquina.

Há dois tipos possíveis:

- a) Relatórios focados em projetos de pesquisa indicados pelo docente
- b) Relatórios focados em outros projetos de pesquisa

Os relatórios deverão ser sucintos com até três páginas (fonte Arial 11 e espaço 1,5) onde devem constar as seguintes seções obrigatórias.

Título

Equipe

Introdução

Métodos

Referências

Requisitos

Seguem as descrições do que se espera em cada seção do relatório. O título deve trazer de forma clara o sistema a ser modelado por técnicas de aprendizado de máquina. Pode ser algo como o seguinte exemplo: Modelos de Aprendizado de Máquina para Prever a Eficácia de Fármacos Anticâncer.

O ideal é que o título não passe de 25 palavras. Na descrição da equipe, o estudante deve ser o primeiro nome. Os outros nomes são facultativos. Na introdução deve constar uma breve apresentação do problema científico a ser tratado (sistema a ser modelado) com as abordagens de aprendizado de máquina. O problema deve se enquadrar como um projeto de aprendizado de máquina, onde espera-se a modelagem computacional de um sistema complexo. O objetivo é que o estudante tenha a visão do que é modelar sistemas a partir de métodos de aprendizado de máquina como regressão, classificação e clustering, mas não necessariamente restritos a estes.

O estudante pode contar com base de dados públicas ou dados de seu grupo de pesquisa. Esses dados serão submetidos ao processo de modelagem computacional por meio de abordagens de aprendizado de máquina. A condição *sine qua non* é que tenhamos disponíveis dados tabuláveis e com variáveis independentes bem definidas e pelo menos uma variável alvo determinada. Ao final da introdução deve constar um

parágrafo único, onde destaca-se o objetivo principal, ou seja, o que se espera conseguir com o uso de abordagens de aprendizado de máquina para modelar o sistema proposto.

Na seção de métodos, os estudantes farão uma breve descrição das técnicas de aprendizado usadas para resolver o problema originalmente proposto. Nesta seção deverá constar o conjunto de métricas propostas para a análise do poder de previsão dos modelos gerados por abordagens de aprendizado de máquina. Espera-se que uso de métricas definidas sejam aplicadas para a avaliação do poder de previsão do modelo de aprendizado de máquina gerado. Há indicações de métricas recentes (Walsh I, Fishman D, Garcia-Gasulla D, Titma T, Pollastri G; ELIXIR Machine Learning Focus Group; Harrow J, Psomopoulos FE, Tosatto SCE. DOME: recommendations for supervised machine learning validation in biology. *Nat Methods*. 2021; 18(10): 1122-1127) e espera-se que os estudantes adotem esses conjuntos de métricas para avaliação do modelo. Para as seções de introdução e métodos, espera-se a citação de referências que devem ser artigos científicos. Os artigos citados devem seguir o seguinte formato.

de Azevedo WF. Application of Machine Learning Techniques for Drug Discovery. *Curr Med Chem*. 2021; 28(38): 7805-7807.

Walsh I, Fishman D, Garcia-Gasulla D, Titma T, Pollastri G; ELIXIR Machine Learning Focus Group; Harrow J, Psomopoulos FE, Tosatto SCE. DOME: recommendations for supervised machine learning validation in biology. *Nat Methods*. 2021; 18(10): 1122-1127.