

Fundamentos de Data Stream Mining

Fabício Enembreck PhD

Programa de Pós-Graduação em Informática da PUCPR (PPGIa)

Especialização em Ciência de Dados

2020

Estrutura do Curso

- Contato: email do BB
- 6 Encontros
 - 09/05, 23/05, 06/06, 20/06, 04/07, 18/07
- Fontes de Informação (disponíveis no BB)
 - Livros
 - Tutoriais
 - Manuais
 - Artigos
 - Teses e Dissertações

Estrutura do Curso

- Avaliação da disciplina
- 2 desafios (25% cada)
 - Atividades desenvolvidas em sala com entrega no final da aula
- 1 projeto (50%)
 - Maior parte desenvolvido fora da aula
 - Teremos um encontro para alinhar andamento e conformidade aos requisitos e outro para defesa/apresentação
- Todas as atividades em grupos de até 4 membros

Estrutura do Curso

- Encontro 1
 - Introdução a DSM
 - Motivação
 - Áreas relacionadas da computação
 - Fundamentos conceituais
 - Técnicas elementares

Estrutura do Curso

- Encontro 2
 - Introdução ao MOA – Massive Online Analysis
 - <https://moa.cms.waikato.ac.nz/>
 - <http://www.jmlr.org/papers/volume11/bifet10a/bifet10a.pdf>
 - Primeiros experimentos
 - Detectores de Mudança
 - Experimentos com detectores

Estrutura do Curso

- Encontro 3
 - Introdução a Classificação em Data Streams
 - Procedimentos de avaliação em Data Streams
 - Experimentos com Classificadores
 - Primeiro Desafio

Estrutura do Curso

- Encontro 4
 - Principais Algoritmos de Ensembles para Classificação em Data Streams
 - Experimentos com Ensembles
 - Segundo Desafio

Estrutura do Curso

- Encontro 5
 - Desenvolvimento de Projeto de DSM
 - Seleção de Dados
 - Seleção de Algoritmos
 - Seleção do Procedimento de Avaliação
 - Seleção de Métricas
 - Geração e comparação de resultados

Estrutura do Curso

- Encontro 6
 - Defesa/Apresentação de Projeto