



O que é? O grupo deverá ser uma integração entre pessoas que buscam aprender e desenvolver técnicas de Machine Learning de forma conjunta. Não é um curso ministrado por um instrutor, mas sim uma junção de esforços para que os interessados aprendam juntos em encontros semanais.

Objetivo: Promover um ambiente de estudo teórico e prático contínuo de técnicas de Machine Learning aplicadas à Engenharia de Recursos Hídricos e Ambiental dentro do PPGERHA.

Coordenador: Prof. Dr. Daniel Henrique Marco Detzel

Objetivos específicos:

1. Desenvolver um ambiente de aprendizado do zero de técnicas de Machine Learning implementadas através da linguagem Python;
2. Fomentar discussões acerca de implementação de algoritmos e aplicações na área de recursos hídricos e ambiental, novas ferramentas e estudos etc.

Público alvo: Alunos e professores do Programa de Pós-graduação em Engenharia de Recursos Hídricos e Ambiental.

Local: PH16 – DHS/UFPR

Encontros: Terças, 17h15 - 19h15

Cronograma 2019/1º semestre

Atividade	Data prevista
Palestra ML – Prof. Prof. Luiz Eduardo Soares de Oliveira	21/03
<i>Início do Módulo I</i>	<i>Abril</i>
Introdução/Princípios do Python	02/04
Pré-processamento de dados	09/04
Regressão Linear Simples	16/04
Regressão Linear Múltipla	23/04
Regressão Linear Polinomial	30/04
Support Vector Regression (SVR)	07/05
Decision Tree Regression	
Random Forest Regression	14/05
Avaliação da Performance de Modelos de Regressão	
Regressão Logística	21/05
K-Nearest Neighbors (KNN)	28/05
<i>Recesso e Matrícula PPGERHA</i>	<i>03/06 – 09/06</i>
Support Vector Machine (SVM)	11/06
Kernel SVM	
Naïve Bayes	18/06
Decision Tree Classification	
Random Forest Classification	25/06
Avaliação da performance dos modelos de classificação	
Agrupamento K-Means	02/07
Agrupamento Hierárquico	
Apriori	09/07
Eclat	
Upper Confidence Bound (UCB)	16/07
Thompson Sampling	
<i>Recesso e Matrícula PPGERHA</i>	<i>22/07 - 03/08</i>
Natural Language Processing	06/08
Redes Neurais Artificiais	
Análise de Componentes Principais	13/08
Análise Discriminante Linear (LDA)	
10.3: Kernel PCA	
<i>Módulo II</i>	<i>20/08 -27/08</i>

Bônus (se der tempo): Seleção de modelo e “boosting”

Programação

Palestra de Boas Vindas: “Machine Learning and the Need for Reproducibility”.

Palestrante: Prof. Luiz Eduardo Soares de Oliveira

Essa palestra terá como objetivo apresentar a área de Machine Learning, suas aplicações, tendências para o futuro e principais técnicas utilizadas.

Módulo I: Machine Learning do zero

Carga-horária: 34 horas

Objetivo: Aprendizado dos modelos mais importantes de aprendizado de máquina, incluindo como criá-los a partir do zero, as principais habilidades em preparação de dados, validação de modelos e criação de produtos de dados.

Material:

- Vídeo-aula da UDEMY (Machine Learning A-Z: Hands on Python) – Kirill Emerenko, Hadelin de Ponteves, e outros materiais gratuitos.
- Templates de código
- Dados .CSV

Método: Exposição teórica (vídeo-aula) e prática de programação em linguagem Python

Conteúdo do Curso

Seção 1: Introdução ao Grupo e Princípios do Python

Seção 2: Pré-processamento de dados

Seção 3: Regressão:

3.1 Regressão Linear Simples

3.2 Regressão Linear Múltipla

3.3 Regressão Polinomial

3.4 Support Vector Regression (SVR)

3.5 Decision Tree Regression

3.6 Random Forest Regression

3.7 Avaliação da Performance de Modelos de Regressão

Seção 4: Classificação

4.1 Regressão Logística

4.2 K-Nearest Neighbors (KNN)

4.3 Support Vector Machine (SVM)

4.4 Kernel SVM

4.5 Naïve Bayes

4.6 Decision Tree Classification

4.7 Random Forest Classification

4.8 Avaliação da performance dos modelos de classificação

Seção 5: Agrupamento

5.1 Agrupamento K-Means

5.2 Agrupamento Hierárquico

Seção 6: Association Rule Learning

6.1 Apriori

6.2 Eclat

Seção 7: Aprendizado por Reforço

7.1 Upper Confidence Bound (UCB)

7.2 Thompson Sampling

Seção 8: Natural Language Processing

Seção 9: Deep Learning

9.1: Redes Neurais Artificiais

9.2 Rede Neural Convolucional

Seção 10: Redução de Dimensionalidade

10.1: Análise de Componentes Principais

10.2: Análise Discriminante Linear (LDA)

10.3: Kernel PCA

Seção 11: Seleção de modelo e “boosting”

11.1: Seleção de modelo

11.2: XGBoost

Módulo II: Implementação prática em projetos de pesquisa

Agosto de 2019

Neste módulo, estaremos aptos a aplicar os conteúdos estudados em nossas áreas de pesquisa, e o grupo funcionará como uma forma de estudo e implementação de algoritmos, discussão e análise de resultados, divulgação de trabalhos e outros objetivos que o grupo achar pertinente.